Perturbation theory for Hamiltonian matrices and the distance to bounded-realness[§]

R. Alam[*] S. Bora[*] M. Karow[†] V. Mehrmann[†] J. Moro [‡]

May 8, 2010

Abstract

Motivated by the analysis of passive control systems, we undertake a detailed perturbation analysis of Hamiltonian matrices that have eigenvalues on the imaginary axis. We construct minimal Hamiltonian perturbations that move and coalesce eigenvalues of opposite sign characteristic to form multiple eigenvalues with mixed sign characteristics, which are then moved from the imaginary axis to specific locations in the complex plane by small Hamiltonian perturbations. We also present a numerical method to compute upper bounds for the minimal perturbations that move all eigenvalues of a given Hamiltonian matrix outside a vertical strip along the imaginary axis.

Keywords. Hamiltonian matrix, Hamiltonian eigenvalue problem, perturbation theory, passive system, bounded-realness, purely imaginary eigenvalues, sign characteristic, Hamiltonian pseudospectra, structured mapping problem, distance to bounded-realness
AMS subject classification. 93B36, 93B40, 49N35, 65F15, 93B52, 93C05.

1 Introduction

In this paper we discuss the perturbation theory for eigenvalues of Hamiltonian matrices and the explicit construction of small perturbations that move eigenvalues from the imaginary axis. With $\mathbb{F}^{k,\ell}$ denoting the vector space of real ($\mathbb{F}=\mathbb{R}$) or complex ($\mathbb{F}=\mathbb{C}$) $k\times\ell$ matrices, a matrix $\mathcal{H}\in\mathbb{F}^{2n,2n}$ is called *Hamiltonian* if $(\mathcal{H}\mathcal{J}_n)^\star=\mathcal{H}\mathcal{J}_n$, where $\mathcal{J}_n=\begin{bmatrix}0 & I_n\\ -I_n & 0\end{bmatrix}$ and I_n is the $n\times n$ identity matrix, (we suppress the subscript n, if the dimension is clear from the context). In order to simplify the presentation and to treat the real and the complex case together, we use * to denote ^{T} in the real case and * in the complex case.

1.1 The distance to bounded-realness

It is well-known [22, 30] that the spectrum of Hamiltonian matrices is symmetric with respect to the imaginary axis, i.e., eigenvalues occur in pairs $(\lambda,-\bar\lambda)$ in the complex case or quadruples

[*]Department of Mathematics, IIT Guwahati, Assam, India; `email: {rafik,shbora}@iitg.ernet.in`

[†]Institut für Mathematik, Ma 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG; `email: {karow,mehrmann}@math.tu-berlin.de`

[‡]Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911-Leganés, Spain `email: jmoro@math.uc3m.es`. Research partially supported by the Spanish Ministerio de Ciencia y Tecnología under grant MTM2006-05361.

[§]Research supported by *Deutsche Forschungsgemeinschaft*, via the DFG Research Center MATHEON in Berlin.

$(\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda})$ in the real case. This eigenvalue symmetry degenerates if there are eigenvalues on the imaginary axis. The existence of purely imaginary eigenvalues typically leads to difficulties for numerical methods in control [5, 30]. If purely imaginary eigenvalues occur, then in some applications (see, e.g., Section 1.2) one perturbs the Hamiltonian matrix in such a way that the eigenvalues are moved away from the imaginary axis. We formulate this as our first problem.

Problem A: *Given a Hamiltonian matrix $\mathcal{H}$ that has purely imaginary eigenvalues, determine (in some norm to be specified) the smallest Hamiltonian perturbation $\Delta\mathcal{H}$ such that for the resulting perturbed matrix $\mathcal{H} + \Delta\mathcal{H}$ an arbitrary small generic Hamiltonian perturbation will move all the eigenvalues off the imaginary axis. Since checking the existence of purely imaginary eigenvalues of a Hamiltonian matrix is used in the context of the Bounded Real Lemma [2] we call this distance* the distance to bounded-realness.

The converse of this problem, i.e., to determine the smallest Hamiltonian perturbation of a Hamiltonian matrix so that all eigenvalues of the resulting perturbed matrix are purely imaginary, i.e., the *distance to non-bounded-realness*, has been recently studied on the basis of so called μ-values and spectral value sets in [18].

While the distance to bounded-realness is an important quantity that has to be determined in order to characterize whether it is possible to find a perturbation that moves all eigenvalues off the imaginary axis, in applications (see, e.g., Section 1.2) often a modified question is more important.

Problem B: *Given a Hamiltonian matrix $\mathcal{H}$ that has purely imaginary eigenvalues, determine (in some norm to be specified) the smallest Hamiltonian perturbation $\Delta\mathcal{H}$ such that the resulting perturbed matrix $\mathcal{H} + \Delta\mathcal{H}$ has all eigenvalues robustly bounded away from the imaginary axis, i.e., all eigenvalues $\tilde{\lambda}$ of $\mathcal{H} + \Delta\mathcal{H}$ lie outside of an open vertical strip $S_\tau = \{z \in \mathbb{C} \mid -\tau < \Re z < \tau\}$ $(\tau \geq 0)$ along the imaginary axis.*

In this paper we discuss numerical procedures for the solution of both Problems **A** and **B**, and obtain upper bounds for the smallest perturbations. We mention that determination of the minimal perturbation is in general a difficult non-convex optimization problem, see [8].

1.2 Passivation

The main motivation for studying the perturbation problems that we have discussed in the previous subsection is the following. Consider a linear time-invariant control system

$$\begin{aligned} \dot{x} &= Ax + Bu, \ x(0) = 0, \\ y &= Cx + Du, \end{aligned} \tag{1}$$

with matrices $A \in \mathbb{F}^{n,n}$, $B \in \mathbb{F}^{n,m}$, $C \in \mathbb{F}^{p,n}$, $D \in \mathbb{F}^{p,m}$. Here u is the input, x the state, and y the output.

Suppose that the homogeneous system is asymptotically stable, i.e., all eigenvalues of A are in the open left half complex plane and that D is square and nonsingular. Then, see e.g., [2], the system is called *passive*, if there exists a nonnegative scalar valued function Θ such that the *dissipation inequality*

$$\Theta(x(t_1)) - \Theta(x(t_0)) \leq \int_{t_0}^{t_1} u^\star y + y^\star u \, dt$$

holds for all $t_1 \geq t_0$, i.e., the system absorbs supply energy.

In real world applications the system model (1) is typically subject to several approximations. Often the real physical problem, e.g., the determination of the electric or magnetic field associated with an electronic device is infinite dimensional and it is approximated by a finite element or finite difference model [17], or the system is nonlinear and the linear model is obtained by a linearization. The system may also be obtained by a realization or system identification [6, 16, 39] or it may be the result of a model reduction procedure [2].

If one uses an approximated model, then it is in general not clear that the property of passivity will be preserved, and typically it is not, i.e., the approximation process makes the passive system non-passive. Since passivity is an important physical property (a passive system does not generate energy), one then approximates the non-passive system by a (hopefully) nearby passive system, by introducing small (minimal) perturbations of A, B, C, D, see [6, 8, 15, 39, 40].

Typically, one has an estimate or even a bound for the approximation error in the original system approximation and then one tries to keep the perturbations within these bounds. So from the application point of view it may not be necessary to really determine the minimal perturbation, a perturbation that stays within the range of the already committed approximation errors is sufficient. But from a system theoretical point of view, it is also interesting to find a value or a bound for the smallest perturbation that makes a non-passive system passive. In general it is an open problem to determine this minimal perturbation explicitly, instead one uses optimization methods, see [6, 8, 9] or *ad hoc* perturbation methods [14, 15, 39, 38], see also [40] for a recent improvement of the method in [15].

The converse problem, i.e., to compute the smallest perturbation that makes a passive system non-passive has recently been studied in [33], again using optimization techniques.

At first sight the passivation problem seems not related to the perturbation problem for Hamiltonian matrices. However, it is well known [2, 15] that one can check whether an asymptotically stable system is passive by checking whether the *Hamiltonian matrix*

$$\mathcal{H} = \begin{bmatrix} F & G \\ H & -F^\star \end{bmatrix} := \begin{bmatrix} A - BR^{-1}C & -BR^{-1}B^\star \\ -C^\star R^{-1}C & -(A - BR^{-1}C)^\star \end{bmatrix} \tag{2}$$

has no purely imaginary eigenvalues, where we have set $R = D + D^\star$. Thus one can use the distance to bounded-realness, i.e., perturbations that solve Problems **A** and **B**, to construct perturbations that make the system passive. This topic will be discussed in forthcoming work.

The paper is organized as follows: In Section 2 we introduce the notation and briefly present some preliminary results. The perturbation theory for eigenvalues, in particular purely imaginary eigenvalues of Hamiltonian matrices is reviewed in Section 3. Hamiltonian perturbations moving purely imaginary eigenvalues of a Hamiltonian matrix to specific points in the complex plane are discussed in Section 4. The minimal perturbations or bounds of minimal perturbations are discussed in Section 5. A numerical method to compute approximate solutions of Problems **A** and **B** for the spectral norm $\|\cdot\|_2$ is discussed in Section 6.

2 Preliminaries

By $\mathbb{C}_+$ and $\mathbb{C}_-$, respectively, we denote the positive right half and negative left half complex plane. For $X \in \mathbb{F}^{n,m}$ of full column rank, we denote by $X^+ := (X^\star X)^{-1}X^\star$ the Moore-Penrose inverse of X, see e.g. [13]. For $A \in \mathbb{F}^{n,n}$, a subspace $\mathcal{X} \subseteq \mathbb{F}^n$ is said to be *A-invariant* if $Ax \in \mathcal{X}$ for any $x \in \mathcal{X}$. In this case we denote by $\Lambda(A|\mathcal{X})$ the spectrum of the restriction of the linear operator A to the subspace $\mathcal{X}$. Let $X \in \mathbb{F}^{n,d}$ be a full column rank matrix such that $\mathcal{X} = \mathrm{range}(X)$. Then $\mathcal{X}$ is A-invariant if $AX = XR$ for some $R \in \mathbb{F}^{d,d}$, and we then have $\Lambda(A|\mathcal{X}) = \Lambda(R)$.

It is well-known [32, 35, 36] that the Hermitian form

$$(x, y) \mapsto i x^\star J y, \quad x, y \in \mathbb{F}^{2n} \tag{3}$$

plays an important role in the perturbation theory of Hamiltonian eigenvalues. If $x^\star J y = 0$, then x and y are said to be J-*orthogonal*. Subspaces $\mathcal{X},\ \mathcal{Y} \subseteq \mathbb{F}^{2n,2n}$ are said to be J-*orthogonal* if $x^\star J y = 0$ for all $x \in \mathcal{X},\ y \in \mathcal{Y}$. A subspace $\mathcal{X} \subseteq \mathbb{F}^{2n,2n}$ is said to be J-*neutral* if $x^\star J x = 0$ for all $x \in \mathcal{X}$. $\mathcal{X}$ is said to be J-*nondegenerate* if for any $x \in \mathcal{X} \setminus \{0\}$ there exists $y \in \mathcal{X}$ such that $x^\star J y \neq 0$.

Nondegenerate invariant subspaces for Hamiltonian matrices are characterized by the following theorem, where for a set of complex numbers $\Xi = \{\xi_1, \ldots, \xi_k\}$ we denote by $\overline{\Xi}$ the set of conjugates of the elements of Ξ.

Theorem 2.1 *Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be invariant subspaces of the Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$. Suppose that $\Lambda(\mathcal{H}|\mathcal{X}_1) \cap \left(-\overline{\Lambda(\mathcal{H}|\mathcal{X}_2)}\right) = \emptyset$. Then $x_1^\star J x_2 = 0$ for all $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$.*

Suppose, additionally, that $\mathcal{X}_1 \oplus \mathcal{X}_2 = \mathbb{F}^{2n}$. Then $\mathcal{X}_1$ and $\mathcal{X}_2$ are J-nongenerate.

Proof. Let $X_k \in \mathbb{F}^{2n,p_k}$ be a matrix whose columns form a basis of $\mathcal{X}_k$, $k = 1, 2$. Then $\mathcal{H} X_k = X_k R_k$, and the matrix $R_k \in \mathbb{F}^{p_k,p_k}$ satisfies $\Lambda(R_k) = \Lambda(\mathcal{H}|\mathcal{X}_k)$. Consider the *Sylvester operator* $\mathcal{S}(Z) = R_1^\star Z + Z R_2$, $Z \in \mathbb{F}^{p_1,p_2}$. We have

$$\begin{aligned}
\mathcal{S}(X_1^\star J X_2) &= R_1^\star X_1^\star J X_2 + X_1^\star J X_2 R_2 \\
&= -(J X_1 R_1)^\star X_2 + X_1^\star (J X_2 R_2) \\
&= -(J \mathcal{H} X_1)^\star X_2 + X_1^\star (J \mathcal{H} X_2) \\
&= -X_1^\star (J\mathcal{H})^\star X_2 + X_1^\star (J\mathcal{H}) X_2 \\
&= 0.
\end{aligned}$$

Furthermore, by assumption $0 \notin \Lambda(R_1^\star) + \Lambda(R_2)$ and, thus, the Sylvester operator $\mathcal{S}$ is nonsingular [24]. Hence, we have $X_1^\star J X_2 = 0$ and this completes the proof of the first claim.

For the second part, suppose that $\mathcal{X}_1 \oplus \mathcal{X}_2 = \mathbb{F}^{2n}$ and that $\mathcal{X}_1$ is degenerate. Then there exists $x_1 \in \mathcal{X}_1 \setminus \{0\}$ such that $x_1^\star J x = 0$ for all $x \in \mathcal{X}_1$. However, we also have $x_1^\star J x = 0$ for all $x \in \mathcal{X}_2$. This yields $x_1^\star J = 0$, contradicting the nonsingularity of J. □

If in Theorem 2.1 we have $\mathcal{X}_1 \oplus \mathcal{X}_2 = \mathbb{F}^{2n}$, and if there exists a basis $\{s_1, \ldots, s_n, s_{n+1}, \ldots, s_{2n}\} \subset \mathbb{F}^{2n}$ such that $\mathrm{span}\{s_1, \ldots, s_n\} = \mathcal{X}_1$ and $\mathrm{span}\{s_{n+1}, \ldots, s_{2n}\} = \mathcal{X}_2$, then the matrix $\mathcal{S} = [s_1, \ldots, s_{2n}]$ is *symplectic*, i.e., $\mathcal{S}^\star J \mathcal{S} = J$. In this case the basis $\{s_1, \ldots, s_n, s_{n+1}, \ldots, s_{2n}\}$ can even be chosen so that $\mathcal{S}$ is *unitary (orthogonal in the real case) and symplectic*, i.e. $\mathcal{S}$ is symplectic and $\mathcal{S}^\star \mathcal{S} = I$, see [3, 27].

We immediately have the following corollary, see e.g. [22].

Corollary 2.2 *Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ be Hamiltonian. Let $i\alpha_1, \ldots, i\alpha_p \in i\mathbb{R}$ be the purely imaginary eigenvalues of $\mathcal{H}$ and let $\lambda_1, \ldots, \lambda_q \in \mathbb{C}$ be the eigenvalues of $\mathcal{H}$ with negative real part. Then the $\mathcal{H}$-invariant subspaces $\ker(\mathcal{H} - i\alpha_k I)^{2n}$ and $\ker(\mathcal{H} - \lambda_j I)^{2n} \oplus \ker(\mathcal{H} + \overline{\lambda}_j I)^{2n}$ are pairwise J-orthogonal. All these subspaces are J-nondegenerate. The subspaces*

$$\begin{aligned}
\mathcal{X}_-(\mathcal{H}) &:= \textstyle\bigoplus_{j=1}^{q} \ker(\mathcal{H} - \lambda_j I)^{2n}, \\
\mathcal{X}_+(\mathcal{H}) &:= \textstyle\bigoplus_{j=1}^{q} \ker(\mathcal{H} + \overline{\lambda_j} I)^{2n}
\end{aligned}$$

are J-neutral.

There are several viewpoints that can be taken to perform the perturbation analysis for Hamiltonian matrices. We will mostly work with the quadratic form (3). However, to set things in perspective, and to introduce some of the necessary terminology we also use the normal and condensed forms for Hamiltonian matrices. For this we need the matrices

$$P_r = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}, \; \hat{P}_r = \begin{bmatrix} & & (-1)^0 \\ & \iddots & \\ (-1)^{r-1} & & \end{bmatrix}, \; N_r = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

and

$$N_r(a) = aI_r + N_r, \quad N_r(a,b) = I_r \otimes \begin{bmatrix} a & b \\ -b & a \end{bmatrix} + N_r \otimes I_2.$$

Theorem 2.3 ([10, 11, 23, 25]) *For a Hamiltonian matrix $\mathcal{H} \in \mathbb{C}^{2n,2n}$ there exists a nonsingular matrix $\mathcal{S}$ such that*

$$\mathcal{S}^{-1}\mathcal{H}\mathcal{S} = \operatorname{diag}(H_1, \ldots, H_m), \quad \mathcal{S}^\star J_n \mathcal{S} = \operatorname{diag}(Z_1, \ldots, Z_m),$$

where each pair (H_j, Z_j) is of one of the following forms:

a) $H_j = iN_{n_j}(\gamma_j)$, $Z_j = is_j P_{n_j}$, where $\gamma_j \in \mathbb{R}$ and $s_j = \pm 1$, corresponding to an $n_j \times n_j$ Jordan block for the purely imaginary eigenvalue $i\gamma_j$.

b) $H_j = \begin{bmatrix} N_{n_j}(\lambda_j) & 0 \\ 0 & -[N_{n_j}(\lambda_j)]^\star \end{bmatrix}$, $Z_j = \begin{bmatrix} 0 & I_{n_j} \\ -I_{n_j} & 0 \end{bmatrix} = J_{n_j}$, where $\lambda_j = a_j + ib_j$ with $a_j, b_j \in \mathbb{R}$ and $a_j \neq 0$, corresponding to an $n_j \times n_j$ Jordan block for each of the eigenvalues λ_j, $-\bar{\lambda}_j$.

The scalars s_j in Theorem 2.3 are called the *sign characteristic* of the pair $(\mathcal{H}, J)$ associated with the purely imaginary eigenvalues, and they satisfy $\sum_{j=1}^{2n} s_j = 0$. The sign characteristics play an important role in the structured perturbation analysis of Hamiltonian matrices see e.g. [12, 22, 32, 34].

In the real case the canonical form is as follows.

Theorem 2.4 ([11, 23, 25, 42]) *Let $\mathcal{H} \in \mathbb{R}^{2n,2n}$ be a Hamiltonian matrix. Then there exists a real nonsingular matrix $\mathcal{S}$ such that*

$$\mathcal{S}^{-1}\mathcal{H}\mathcal{S} = \operatorname{diag}(H_1, \ldots, H_m), \quad \mathcal{S}^T J_n \mathcal{S} = \operatorname{diag}(Z_1, \ldots, Z_m),$$

where each pair (H_j, Z_j) is of one of the following forms:

(a.1) $H_j = t_j \begin{bmatrix} 0 & (-1)^0 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & (-1)^{2n_j-2} \\ & & & 0 \end{bmatrix}$, $Z_j = \hat{P}_{2n_j}$ where $t_j = \pm 1$, corresponding to a $2n_j \times 2n_j$ Jordan block for the eigenvalue 0,

(a.2) $H_j = \begin{bmatrix} N_{2n_j+1} & 0 \\ 0 & -N_{2n_j+1}^T \end{bmatrix}$, $Z_j = \begin{bmatrix} 0 & I_{2n_j+1} \\ -I_{2n_j+1} & 0 \end{bmatrix} = J_{2n_j+1}$, *corresponding to two* $(2n_j+1) \times (2n_j+1)$ *Jordan blocks for the eigenvalue* 0.

b) $H_j = \begin{bmatrix} 0 & N_{n_j}(\gamma_j) \\ -N_{n_j}(\gamma_j) & 0 \end{bmatrix}$, $Z_j = s_j \begin{bmatrix} 0 & P_{n_j} \\ -P_{n_j} & 0 \end{bmatrix}$, *where* $0 < \gamma_j \in \mathbb{R}$ *and* $s_j = \pm 1$, *corresponding to an* $n_j \times n_j$ *Jordan block for each of the purely imaginary eigenvalues* $\pm i\gamma_j$.

c) $H_j = \begin{bmatrix} N_{n_j}(\beta_j) & 0 \\ 0 & -[N_{n_j}(\beta_j)]^T \end{bmatrix}$, $Z_j = \begin{bmatrix} 0 & I_{n_j} \\ -I_{n_j} & 0 \end{bmatrix} = J_{n_j}$ *where* $0 < \beta_j \in \mathbb{R}$, *corresponding to an* $n_j \times n_j$ *Jordan block for each of the real eigenvalues* β_j *and* $-\beta_j$.

d) $H_j = \begin{bmatrix} N_{n_j}(a_j, b_i) & 0 \\ 0 & -[N_{n_j}(a_j, b_j)]^T \end{bmatrix}$, $Z_j = \begin{bmatrix} 0 & I_{2n_j} \\ -I_{2n_j} & 0 \end{bmatrix} = J_{2n_j}$ *where* $0 < a_j, b_j \in \mathbb{R}$, *corresponding to an* $n_j \times n_j$ *Jordan block for each of the eigenvalues* $a_j + ib_j$, $-a_j + ib_j$, $a_j - ib_j$ *and* $-a_j - ib_j$.

It should be noted that in the real case we have two sets of sign characteristics $\{t_j\}$, $\{s_j\}$ for the pair $(\mathcal{H}, J)$.

Note further that both in the real and complex case, an analogous canonical form can be constructed where the transformation matrix is symplectic, see [25]. The normal form under symplectic transformations forms the basis for the computation of eigenvalues, sign characteristics, eigenvectors and invariant subspaces of Hamiltonian matrices. But since the group of symplectic matrices is not compact, to obtain backward stable numerical methods it is important to use unitary (orthogonal) symplectic matrices for the transformations. In this case, in general, we cannot get the complete spectral information but only a condensed form, the (partial) Hamiltonian Schur form.

Lemma 2.5 *[26, 30] Given a Hamiltonian matrix* $\mathcal{H} \in \mathbb{F}^{2n,2n}$, *there exist a unitary symplectic (real orthogonal symplectic if* $\mathbb{F} = \mathbb{R}$*) matrix* $Q \in \mathbb{F}^{2n,2n}$ *such that*

$$\mathcal{H}_1 = Q^\star \mathcal{H} Q = \left[\begin{array}{cc|cc} F_{11} & F_{12} & G_{11} & G_{12} \\ 0 & F_{22} & G_{21} & G_{22} \\ \hline 0 & 0 & -F_{11}^\star & 0 \\ 0 & H_{22} & -F_{12}^\star & -F_{22}^\star \end{array}\right], \tag{4}$$

where F_{11} *is upper triangular (quasi-upper triangular in the real case) and has only eigenvalues in the open left half plane, while the submatrix*

$$\left[\begin{array}{c|c} F_{22} & G_{22} \\ \hline H_{22} & -F_{22}^\star \end{array}\right],$$

has only purely imaginary eigenvalues.

If there are no purely imaginary eigenvalues, then this latter block is void, and this becomes a Hamiltonian Schur form.

Under further conditions, see [7, 25, 26] a Hamiltonian Schur form also exists if purely imaginary eigenvalues occur.

We now discuss the perturbation theory for purely imaginary eigenvalues of Hamiltonian matrices.

3 Perturbation theory for Hamiltonian matrices

In this section we discuss perturbation results for Hamiltonian matrices. In particular, we analyze how purely imaginary eigenvalues of Hamiltonian matrices behave under Hamiltonian perturbations and then we characterize when small perturbations allow to move purely imaginary eigenvalues away from the imaginary axis, see also [21, 32, 34, 35, 36]). To be more precise, given a Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$ with a purely imaginary eigenvalue $i\alpha$, our primary aim is to determine a minimal Hamiltonian perturbation $\Delta\mathcal{H}$ such that $i\alpha$ moves away from the imaginary axis to some specified location in the complex plane, when $\mathcal{H}$ is perturbed to $\mathcal{H} + \Delta\mathcal{H}$. By minimal perturbation we mean that $\Delta\mathcal{H}$ has the smallest norm, either in the Frobenius or in the spectral norm.

It is well-known that the spectral perturbation theory for Hamiltonian matrices [32, 34, 35], in particular for the purely imaginary eigenvalues, is substantially different from the well-known classical perturbation theory for eigenvalues and eigenvectors of unstructured matrices, see e.g. [41]. This is demonstrated by the following example, see e.g. [32].

Example 3.1 The Hamiltonian matrices

$$\mathcal{H}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad \mathcal{H}_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

both have a pair of purely imaginary eigenvalues $\pm i$ with algebraic multiplicity 2.

For $\mathcal{H}_1$, where i is a double eigenvalue with opposite sign characteristic, a small Hamiltonian perturbation will generically move the eigenvalue i off the imaginary axis [28, 29], while for $\mathcal{H}_2$, where the sign characteristic of both eigenvalues is the same, only a large Hamiltonian perturbation can achieve this.

Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ be Hamiltonian and suppose that $i\alpha$ is a purely imaginary eigenvalue of $\mathcal{H}$. Let X be a full column rank matrix so that the columns of X span the right invariant subspace $\ker(\mathcal{H} - i\alpha I)^{2n}$ associated with $i\alpha$, i.e.,

$$\mathcal{H}X = XR, \tag{5}$$

where $\Lambda(R) = \{i\alpha\}$. By using the fact that $\mathcal{H}$ is Hamiltonian, we also have

$$X^\star J\mathcal{H} = -R^\star X^\star J. \tag{6}$$

Since also $\Lambda(-R^\star) = \{i\alpha\}$, it follows that the columns of the full column rank matrix $J^\star X$ span the left invariant subspace associated with $i\alpha$. Hence, $(J^\star X)^\star X = X^\star JX$ is nonsingular and the matrix

$$Z = iX^\star JX \tag{7}$$

associated with the Hermitian form (3) is nonsingular. Based on this observation in [32] the following results for the spectral norm $\|\cdot\|_2$ was shown.

Theorem 3.2 *[32] Consider a Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$ with a purely imaginary eigenvalue $i\alpha$ of algebraic multiplicity p. Suppose that $X \in \mathbb{F}^{2n,p}$ satisfies $\operatorname{rank} X = p$ and (5), and that Z as defined in (7) is congruent to $\left[\begin{smallmatrix} I_\pi & 0 \\ 0 & -I_\mu \end{smallmatrix}\right]$ (with $\pi + \mu = p$).*

If $\Delta\mathcal{H}$ is Hamiltonian and $\|\Delta\mathcal{H}\|_2$ is sufficiently small, then $\mathcal{H}+\Delta\mathcal{H}$ has p eigenvalues $\lambda_1,\ldots,\lambda_p$ (counting multiplicity) in the neighborhood of $i\alpha$, among which at least $|\pi-\mu|$ eigenvalues are purely imaginary. In particular, we have the following cases.

1. *If Z is definite, i.e. either $\pi=0$ or $\mu=0$, then all $\lambda_1,\ldots,\lambda_p$ are purely imaginary with equal algebraic and geometric multiplicity, and satisfy*

$$\lambda_j = i(\alpha+\delta_j)+O(\|\Delta\mathcal{H}\|_2^2),$$

where $\delta_1,\ldots,\delta_p$ are the real eigenvalues of the pencil $\lambda Z - X^\star(J\Delta\mathcal{H})X$.

2. *If there exists a Jordan block associated with $i\alpha$ of size larger than 2, then generically for a given $\Delta\mathcal{H}$ some eigenvalues of $\mathcal{H}+\Delta\mathcal{H}$ will no longer be purely imaginary.*

   If there exists a Jordan block associated with $i\alpha$ of size 2, then for any $\epsilon>0$, there always exists a Hamiltonian perturbation matrix $\Delta\mathcal{H}$ with $\|\Delta\mathcal{H}\|_2=\epsilon$ such that some eigenvalues of $\mathcal{H}+\Delta\mathcal{H}$ will have nonzero real part.

3. *If $i\alpha$ has equal algebraic and geometric multiplicity and Z is indefinite, then for any $\epsilon>0$, there always exists a Hamiltonian perturbation matrix $\Delta\mathcal{H}$ with $\|\Delta\mathcal{H}\|_2=\epsilon$ such that some eigenvalues of $\mathcal{H}+\Delta\mathcal{H}$ will have nonzero real part.*

We now revisit the perturbation results in Theorem 3.2 and present them in a form that we can directly use in the construction of small perturbations. In what follows, we show that an imaginary eigenvalue of $\mathcal{H}$ can be moved off the imaginary axis by an arbitrary small Hamiltonian perturbation if and only if $\mathcal{H}$ has a J-neutral eigenvector corresponding to the imaginary eigenvalue. We then describe how to construct such a Hamiltonian perturbation.

Suppose that we wish to construct a Hamiltonian perturbation matrix $\mathcal{E}$ of smallest norm such that an eigenvalue of $\mathcal{H}$ moves to μ, when $\mathcal{H}$ is perturbed to $\mathcal{H}+\mathcal{E}$. If we have constructed such a perturbation, then there exists a vector u such that $(\mathcal{H}+\mathcal{E})u=\mu u$. This means that $\mathcal{E}u=\mu u-\mathcal{H}u=r$. Thus, the resulting $\mathcal{E}$ is a solution of the following *structured mapping problem*, see [1]. *Given $x,b\in\mathbb{F}^{2n}$ find a Hamiltonian matrix $\mathcal{G}$ of smallest norm $\|\mathcal{G}\|$ such that $\mathcal{G}x=b$.* Here $\|\cdot\|$ is either the spectral norm or the Frobenius norm.

To solve this problem in a more general framework, for $X\in\mathbb{F}^{2n,p}$ and $B\in\mathbb{F}^{2n,p}$, we introduce

$$\eta(X,B) := \inf\{\|\mathcal{H}\| : \mathcal{H}\in\mathbb{F}^{2n,2n}, (J\mathcal{H})^\star = J\mathcal{H} \text{ and } \mathcal{H}X=B\}, \tag{8}$$

denoting $\eta(X,B)$ by $\eta_F(X,B)$ for the Frobenius norm and by $\eta_2(X,B)$ for the spectral norm. The following result of [1] provides a complete solution of the Hamiltonian structured mapping problem.

Theorem 3.3 *[1]*

a) 1. *Let $x\in\mathbb{F}^{2n}$ and $b\in\mathbb{F}^{2n}$. Then there exists a Hamiltonian matrix $\mathcal{H}\in\mathbb{F}^{2n,2n}$ such that $\mathcal{H}x=b$ if and only if $x^\star Jb\in\mathbb{R}$.*

2. *If $x^\star Jb$ is real, then*

$$\begin{aligned}\eta_F(x,b) &= \sqrt{2\|b\|_2^2/\|x\|_2^2 - |x^\star Jb|^2/\|x\|_2^4}\\ \eta_2(x,b) &= \|b\|_2/\|x\|_2.\end{aligned}$$

Furthermore, the matrix

$$\mathcal{G}(x,b) := \frac{bx^\star + Jxb^\star J}{\|x\|_2^2} + \frac{(x^\star Jb)\, Jxx^\star}{\|x\|_2^4}$$

is the unique Hamiltonian matrix such that $\mathcal{G}(x,b)x = b$ *and* $\|\mathcal{G}(x,b)\|_F = \eta_F(x,b)$.

3. *If* $\|x\|_2\|b\|_2 \neq |x^\star Jb|$, *then form the Hamiltonian matrix*

$$\mathcal{F}(x,b) := \mathcal{G}(x,b) - \frac{x^\star Jb}{\|b\|_2^2\|x\|_2^2 - |x^\star Jb|^2} J(I - \frac{xx^\star}{x^\star x}) Jbb^\star J(I - \frac{xx^\star}{x^\star x}),$$

otherwise, set $\mathcal{F}(x,b) := \mathcal{G}(x,b)$. *Then* $\mathcal{F}(x,b)x = b$ *and* $\|\mathcal{F}(x,b)\|_2 = \eta_2(x,b)$.

b) 1. *Let* $B \in \mathbb{F}^{2n,p}$ *and* $X \in \mathbb{F}^{2n,p}$. *Suppose that* $\operatorname{rank} X = p$. *Then there exists a Hamiltonian matrix* $\mathcal{H} \in \mathbb{F}^{2n,2n}$ *such that* $\mathcal{H}X = B$ *if and only if* $X^\star JB$ *is Hermitian.*

2. *If* $X^\star JB$ *is Hermitian, then*

$$\begin{aligned}\eta_2(X,B) &= \|B(X^\star X)^{-1/2}\|_2 \\ \eta_F(X,B) &= \sqrt{2\|B(X^\star X)^{-1/2}\|_F^2 - \|(X^\star X)^{-1/2}X^\star JB(X^\star X)^{-1/2}\|_F^2}.\end{aligned}$$

3. *The matrix*

$$\mathcal{G}(X,B) := BX^+ + J(X^+)^\star B^\star J + JXX^+JBX^+, \tag{9}$$

is the unique Hamiltonian matrix such that $\mathcal{G}(X,B)X = B$ *and* $\|\mathcal{G}(X,B)\|_F = \eta_F(X,B)$.

4. *Set* $Z := (X^\star X)^{-1/2}X^\star JB(X^\star X)^{-1/2}$ *and* $\rho := \eta_2(X,B)$. *If* $\rho^2 I - Z^2$ *is nonsingular, then consider the Hamiltonian matrix*

$$\mathcal{F}(X,B) := \mathcal{G}(X,B) + J(I - XX^+)KZK^\star(I - XX^+),$$

where $K := JB(X^\star X)^{-1/2}(\rho^2 I - Z^2)^{-1/2}$. *Then* $\mathcal{F}(X,B)$ *is a Hamiltonian matrix such that* $\mathcal{F}(X,B)X = B$ *and* $\|\mathcal{F}(X,B)\|_2 = \eta_2(X,B)$.

In order to construct a *real* Hamiltonian matrix $\mathcal{H}$ satisfying $\mathcal{H}X = B$ we need the following lemma.

Lemma 3.4 *Let* $A, B \in \mathbb{C}^{n,p}$. *Then* $[A\ \bar{A}][B\ \bar{B}]^+$ *is a real matrix.*

Proof. Let $P = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \in \mathbb{R}^{2p,2p}$. Then $[A\ \bar{A}]P = [\bar{A}\ A]$. Since $P^{-1} = P^\star = P$ we have $P[B\ \bar{B}]^+ = ([B\ \bar{B}]P)^+ = [\bar{B}\ B]^+$. Hence $\overline{[A\ \bar{A}][B\ \bar{B}]^+} = \overline{[A\ \bar{A}]P^2[B\ \bar{B}]^+} = [A\ \bar{A}][B\ \bar{B}]^+$. □

We then have the following minimal real perturbations.

Corollary 3.5 *Let* $B \in \mathbb{C}^{2n,p}$, $X \in \mathbb{C}^{2n,p}$ *and suppose that* $\operatorname{rank}[X\ \bar{X}] = 2p$. *Then there exists a* real *Hamiltonian matrix* $\mathcal{H} \in \mathbb{R}^{2n,2n}$ *such that* $\mathcal{H}X = B$ *if and only if* $X^\star JB$ *is Hermitian and* $X^\star J\bar{B}$ *is symmetric, i.e.,* $(X^\star J\bar{B})^\top = X^\star J\bar{B}$.

If the latter two conditions are satisfied, then with $\mathcal{G}$ *as defined in (9), the matrix* $\mathcal{G}_\mathbb{R} := \mathcal{G}([X\ \bar{X}], [B\ \bar{B}])$ *is a real Hamiltonian matrix with* $\mathcal{G}_\mathbb{R}X = B$. *Furthermore, among all real Hamiltonian matrices* $\mathcal{H}$ *with* $\mathcal{H}X = B$ *the matrix* $\mathcal{G}_\mathbb{R}$ *has the smallest Frobenius norm.*

Proof. If $\mathcal{H}$ is any real matrix with $\mathcal{H}X = B$ then also $\mathcal{H}\bar{X} = \bar{B}$. Hence $\mathcal{H}[X\ \bar{X}] = [B\ \bar{B}]$. By Theorem 3.3 a Hamiltonian matrix $\mathcal{H}$ satisfying this relation exists if and only if $[X\ \bar{X}]^\star J[B\ \bar{B}] =: Z$ is Hermitian. It is easily verified that Z is Hermitian if and only if $X^\star JB$ is Hermitian and $X^\star J\bar{B}$ is symmetric. If these conditions are satisfied then by Theorem 3.3 the matrix $\mathcal{G}_\mathbb{R}$ is Hamiltonian and $\mathcal{G}_\mathbb{R}[X\ \bar{X}] = [B\ \bar{B}]$. Moreover, among all Hamiltonian matrices $\mathcal{H}$ with $\mathcal{H}[X\ \bar{X}] = [B\ \bar{B}]$ the matrix $\mathcal{G}_\mathbb{R}$ has smallest Frobenius norm. The realness of $\mathcal{G}_\mathbb{R}$ follows from Lemma 3.4. $\square$

In this section, we have discussed the structured mapping theorem for Hamiltonian matrices and used it to construct solutions of minimal Frobenius and spectral norm. In the next section we use these results to construct Hamiltonian perturbations that move eigenvalues away from the imaginary axis.

4 Moving eigenvalues by small perturbations

We now discuss in detail how to move an eigenvalue (resp., a group of eigenvalues) of a Hamiltonian matrix by a small Hamiltonian perturbation to a specific location (resp., locations) in the complex plane. We construct Hamiltonian perturbations under the assumption that a J-neutral eigenvector (resp., J-neutral invariant subspace) exists corresponding the eigenvalue (resp., group of eigenvalues).

Theorem 4.1 *Let σ be a set of eigenvalues of a Hamiltonian matrix $\mathcal{H} \in \mathbb{C}^{2n,2n}$ and $X \in \mathbb{C}^{2n,d}$ be a full column rank matrix such that $X^\star JX = 0$ and $\mathcal{H}X = XR$ for some $R \in \mathbb{C}^{d,d}$ with $\Lambda(R) = \sigma$. Then for any $D \in \mathbb{C}^{d,d}$, the matrix $\mathcal{E} = \mathcal{G}(X, XD)$, where $\mathcal{G}(\cdot,\cdot)$ is defined by (9), has the following properties.*

i) The matrix $\mathcal{E}$ is Hamiltonian and satisfies $\mathcal{E} = XDX^+ + J(X^+)^\star D^\star X^\star J$, $\mathcal{E}X = XD$, $\|\mathcal{E}\|_2 = \|XD(X^\star X)^{-1/2}\|_2$ and $\|\mathcal{E}\|_F = \sqrt{2}\,\|\mathcal{E}\|_2$. Further,

$$(\mathcal{H} + t\mathcal{E})X = X(R + tD), \tag{10}$$

for all $t \in \mathbb{R}$, i.e., $\Lambda(R + tD) \subset \Lambda(\mathcal{H} + t\mathcal{E})$ for all $t \in \mathbb{R}$.

ii) When $\mathcal{H}$ is real and the eigenvalues in σ are real, then the matrix X can be chosen to be real, so that $\mathcal{E}$ is a real Hamiltonian matrix provided that D is real.

iii) When $\mathcal{H}$ is real and $\sigma \cap \bar{\sigma} = \emptyset$, then the matrix $\mathcal{K} = \mathcal{G}([X\ \bar{X}], [XD\ \overline{XD}])$ is real Hamiltonian and satisfies $\mathcal{K}X = XD$, where $\bar{\sigma} := \{\bar{\lambda} : \lambda \in \sigma\}$. Further, for all $t \in \mathbb{R}$ we have

$$(\mathcal{H} + t\mathcal{K})X = X(R + tD), \tag{11}$$

i.e., $\Lambda(R + tD) \subset \Lambda(\mathcal{H} + t\mathcal{K})$ for all $t \in \mathbb{R}$.

Proof. Since $X^\star J(XD) = 0$ is Hermitian, by Theorem 3.3, $\mathcal{E}$ is a well defined Hamiltonian matrix, $\mathcal{E} = XDX^+ + J(X^+)^\star D^\star X^\star J$, $\mathcal{E}X = XD$, $\|\mathcal{E}\|_2 = \|XD(X^\star X)^{-1/2}\|_2$ and $\|\mathcal{E}\|_F = \sqrt{2}\,\|\mathcal{E}\|_2$. This proves i).

The assertion in ii) is obvious. So, suppose that $\mathcal{H}$ is real and that $\sigma \cap \bar{\sigma} = \emptyset$. Then we have $\mathcal{H}X = XR$ and $\mathcal{H}\bar{X} = \bar{X}\,\bar{R}$ with $\Lambda(R) \cap \Lambda(\bar{R}) = \emptyset$. Hence $\operatorname{rank}[X, \bar{X}] = 2d$ and by Theorem 2.1 the spaces spanned by the columns of X and $\bar{X}$ are J-orthogonal. Thus,

$X^\star J\bar{X} = 0$. Since $X^\star JXD = 0$ is Hermitian and $X^\star J\overline{XD} = 0$ is symmetric, by Corollary 3.5, the matrix $\mathcal{K}$ is real and Hamiltonian with $\mathcal{K}X = XD$. This proves iii). $\square$

Theorem 4.1 shows that an eigenvalue (resp., a group of eigenvalues) of a Hamiltonian matrix $\mathcal{H}$ can be moved by a small Hamiltonian perturbation if the eigenvalue (resp., group of eigenvalues) is associated with a J-neutral eigenvector (resp., J-neutral $\mathcal{H}$-invariant subspace).

Remark 4.2 If $\lambda \in \mathbb{C}\backslash i\mathbb{R}$ is a non-imaginary eigenvalue of $\mathcal{H}$ and v is an associated eigenvector then v is J-neutral, that is, $v^\star Jv = 0$. Thus by Theorem 4.1, a non-imaginary eigenvalue of $\mathcal{H}$ can be moved in any direction in the complex plane by a small Hamiltonian perturbation. More generally, let σ be a set of eigenvalues of $\mathcal{H}$ such that $\sigma \subset \mathbb{C}_-$ (or equivalently $\sigma \subset \mathbb{C}_+$). Then by Corollary 2.2, there is a full column rank matrix X such that $X^\star JX = 0$ and $\mathcal{H}X = XR$ with $\Lambda(R) = \sigma$, for some matrix R. Hence by Theorem 4.1, the group of eigenvalues σ can be moved en block by a small Hamiltonian perturbation. Moreover, when $\mathcal{H}$ is real and $\sigma \cap \bar{\sigma} = \emptyset$, then the Hamiltonian perturbation can be chosen to be real.

In view of Remark 4.2 we conclude that a non-imaginary eigenvalue (that is, an eigenvalue with nonzero real part) of a Hamiltonian matrix can be moved in any direction in the complex plane by a small Hamiltonian perturbation. However, this property does not hold in the same generality for purely imaginary eigenvalues. Indeed, suppose that $i\alpha$ is an imaginary eigenvalue of $\mathcal{H}$ and v is an associated eigenvector, that is, $\mathcal{H}v = i\alpha v$. Then by the Hamiltonian eigenvalue symmetry, Jv is a left eigenvector of $\mathcal{H}$ corresponding to $i\alpha$, that is, $(Jv)^\star\mathcal{H} = i\alpha(Jv)^\star$. Thus if v is J-neutral then $(Jv)^\star v = -v^\star Jv = 0$. Hence it follows that the algebraic multiplicity of $i\alpha$ must be at least 2. However, the algebraic multiplicity being at least 2 is not enough to remove an imaginary eigenvalue from the imaginary axis by a small Hamiltonian perturbation. The crux of the matter is that the existence of a J-neutral eigenvector is both necessary and sufficient condition for moving an eigenvalue (imaginary or not) of a Hamiltonian matrix in any direction in the complex plane by a small Hamiltonian perturbation. More generally, we have the following result.

Theorem 4.3 *Let $\sigma := \{\lambda_1, \ldots, \lambda_m\}$ be a set of eigenvalues of a Hamiltonian matrix $\mathcal{H} \in \mathbb{C}^{2n,2n}$. Then there exists a Hamiltonian matrix $\mathcal{E}$ such that $\mathcal{H}(t) := \mathcal{H} + t\mathcal{E}$ has a p-dimensional $\mathcal{H}(t)$-invariant subspace $\mathcal{X}(t)$ with $\sigma(t) := \Lambda(\mathcal{H}(t)|\mathcal{X}(t)) \subset \mathbb{C}_-$ (or equivalently $\sigma(t) \subset \mathbb{C}_+$) for $0 < t \leq 1$ and $\sigma(t) \to \sigma$ as $t \to 0$ if and only if the generalized eigenspace $\bigoplus_{k=1}^m \ker(\mathcal{H} - \lambda_k I)^{2n}$ contains a p-dimensional J-neutral $\mathcal{H}$-invariant subspace $\mathcal{X}$ with $\Lambda(\mathcal{H}|\mathcal{X}) = \sigma$.*

Proof. Suppose that $\mathcal{H}X = XR$ with $\Lambda(R) = \sigma$ and $X^\star JX = 0$, where $X \in \mathbb{C}^{2n,p}$ is a full column rank matrix. Then the desired result follows from Theorem 4.1.

Conversely, suppose that there exists a Hamiltonian matrix $\mathcal{E}$ such that $\mathcal{H}(t) := \mathcal{H} + t\mathcal{E}$ has a p-dimensional $\mathcal{H}(t)$-invariant subspace $\mathcal{X}(t)$ with $\sigma(t) := \Lambda(\mathcal{H}(t)|\mathcal{X}(t)) \subset \mathbb{C}_-$ for $0 < t \leq 1$ and $\sigma(t) \to \sigma$ as $t \to 0$. Let $X(t) \in \mathbb{C}^{2n,p}$ be a matrix with orthonormal columns such that $\mathrm{span}(X(t)) = \mathcal{X}(t)$. Then $\mathcal{H}(t)X(t) = X(t)R(t)$ for some $R(t)$ with $\Lambda(R(t)) = \sigma(t)$. By multiplying the former equation from the left with $X(t)^\star$, it follows that $R(t) = X(t)^\star\mathcal{H}(t)X(t)$. Since for $t > 0$, the set $\sigma(t)$ contains no purely imaginary eigenvalue of $\mathcal{H}(t)$, the invariant subspace $\mathcal{X}(t)$ is J-neutral by Corollary 2.2. Thus $X(t)^\star JX(t) = 0$ for $t > 0$. Since the set of $2n$-by-p matrices with orthonormal columns is compact, the limit $X = \lim_{k\to\infty} X(t_k)$ exists for some sequence $\{t_k\}$ with $t_k \to 0$. By continuity, it follows that

$X^\star JX = 0$ and $\mathcal{H}X = XR$, where $R = \lim R(t_k)$. Furthermore, $\Lambda(R) = \lim \sigma(t_k) = \sigma$. Hence $\mathcal{X} := \operatorname{span}(X)$ is a J-neutral $\mathcal{H}$-invariant p-dimensional subspace of $\bigoplus_{k=1}^m \ker(\mathcal{H} - \lambda_k I)^{2n}$ with $\Lambda(\mathcal{H}|\mathcal{X}) = \sigma$. □

A J-neutral $\mathcal{H}$-invariant subspace of $\ker(\mathcal{H} - i\alpha I)^{2n}$ of maximal dimension can be constructed from the canonical form displayed in Theorem 2.4. We, however, will not discuss the construction here because such a subspace is not needed for our purpose.

Corollary 4.4 *An eigenvalue λ of a Hamiltonian matrix $\mathcal{H}$ can be removed from the imaginary axis by an arbitrarily small Hamiltonian perturbation if and only if $\mathcal{H}$ has a J-neutral eigenvector corresponding to λ.*

Note that an imaginary eigenvalue of a Hamiltonian matrix may or may not have a J-neutral eigenvector associated with it. So, if an imaginary eigenvalue does not have a J-neutral eigenvector associated with it then in such a case the eigenvalue cannot be removed from the imaginary axis by a small Hamiltonian perturbation, it has to be combined with another eigenvalue of opposite sign characteristic. In our algorithmic construction we remove one imaginary eigenvalue at a time. Therefore, we first briefly discuss the removal from the imaginary axis of an imaginary eigenvalue by a Hamiltonian perturbation under the assumption that a J-neutral eigenvector exists and then we discuss how to achieve this property. We have the following result which follows from Theorem 4.1.

Theorem 4.5 *Let $i\alpha$ be an imaginary eigenvalue of a Hamiltonian matrix $\mathcal{H} \in \mathbb{C}^{2n,2n}$. Let v be a normalized and J-neutral eigenvector of $\mathcal{H}$ corresponding $i\alpha$, i.e., $\|v\|_2 = 1$, $v^\star Jv = 0$ and $\mathcal{H}v = i\alpha v$. For any $\mu \in \mathbb{C}$, consider the matrices*

$$\mathcal{E}_\mu = \mathcal{G}(v, \mu v) \qquad \text{and} \qquad \mathcal{K}_\mu = \mathcal{G}([v\ \bar{v}], [\mu v\ \overline{\mu v}]),$$

where $\mathcal{G}(\cdot,\cdot)$ is defined by (9). Then $\mathcal{E}_\mu$ and $\mathcal{K}_\mu$ have the following properties.

i) The matrix $\mathcal{E}_\mu$ is Hamiltonian and satisfies $\mathcal{E}_\mu = \mu vv^\star + \bar{\mu}Jvv^\star J$, $\|\mathcal{E}_\mu\|_2 = |\mu|$ and $\|\mathcal{E}_\mu\|_F = \sqrt{2}|\mu|$. Furthermore, $(\mathcal{H} + t\mathcal{E}_\mu)v = (i\alpha + t\mu)v$ for all $t \in \mathbb{R}$, i.e., $i\alpha + t\mu \in \Lambda(\mathcal{H} + t\mathcal{E}_\mu)$ for all $t \in \mathbb{R}$.

ii) If $\mathcal{H}$ is a real matrix and $\alpha = 0$, then the vector v can be chosen to be real in which case $\mathcal{E}_\mu$ is real for all $\mu \in \mathbb{R}$.

iii) Suppose that $\mathcal{H}$ is a real matrix and $\alpha \neq 0$. Then $\mathcal{K}_\mu$ is a real Hamiltonian matrix satisfying $(\mathcal{H} + t\mathcal{K}_\mu)v = (i\alpha + t\mu)v$ and $(\mathcal{H} + t\mathcal{K}_\mu)\bar{v} = (-i\alpha + t\bar{\mu})\bar{v}$. Hence $\{i\alpha + t\mu, -i\alpha + t\bar{\mu}\} \subset \Lambda(\mathcal{H} + t\mathcal{K}_\mu)$ for all $t \in \mathbb{R}$.

The perturbations $\mathcal{E}_\mu$ and $\mathcal{K}_\mu$ constructed in Theorem 4.5 move the imaginary eigenvalue $i\alpha$ away from the imaginary axis. Note, however, that these perturbations may also move the other eigenvalues of $\mathcal{H}$ to unspecified positions. For our algorithmic construction, it is desirable to move eigenvalues one-by-one without affecting the other eigenvalues. The following result provides Hamiltonian perturbations which move only the eigenvalue $i\alpha$ and leave the other eigenvalues unchanged.

Theorem 4.6 *Let $i\alpha$ be an imaginary eigenvalue of a Hamiltonian matrix $\mathcal{H} \in \mathbb{C}^{2n,2n}$. Let v be a normalized and J-neutral eigenvector of $\mathcal{H}$ corresponding $i\alpha$, i.e., $\|v\|_2 = 1$, $v^\star Jv = 0$*

and $\mathcal{H}v = i\alpha v$. Let $w \in \ker(\mathcal{H} - i\alpha I)^{2n}$ be such that $w^\star Jv = 1$. For any $\mu \in \mathbb{C}$, consider the matrices

$$\hat{\mathcal{E}}_\mu = (\mu v w^\star + \bar{\mu} w v^\star) J \qquad \text{and} \qquad \hat{\mathcal{K}}_\mu = \hat{\mathcal{E}}_\mu + \overline{\hat{\mathcal{E}}_\mu}.$$

Then $\hat{\mathcal{E}}_\mu$ and $\hat{\mathcal{K}}_\mu$ have the following properties.

i) The matrix $\hat{\mathcal{E}}_\mu$ is Hamiltonian and $(\mathcal{H} + t\hat{\mathcal{E}}_\mu)v = (i\alpha + t\mu)v$ for all $t \in \mathbb{R}$. Furthermore, $(\mathcal{H} + t\hat{\mathcal{E}}_\mu)x = \mathcal{H}x$ for any $x \in \ker(\mathcal{H} - \lambda I)^{2n}$ and $\lambda \in \Lambda(\mathcal{H}) \setminus \{i\alpha\}$.

ii) Suppose that $\mathcal{H}$ is a real matrix and $\alpha = 0$. Then the vectors v and w can be chosen to be real in which case $\hat{\mathcal{E}}_\mu$ is real for all $\mu \in \mathbb{R}$.

iii) Suppose that $\mathcal{H}$ is a real matrix and $\alpha \neq 0$. Then the matrix $\hat{\mathcal{K}}_\mu$ is a real Hamiltonian matrix satisfying $(\mathcal{H}+t\hat{\mathcal{K}}_\mu)v = (i\alpha+t\lambda)v$, $(\mathcal{H}+t\hat{\mathcal{K}}_\mu)\bar{v} = (-i\alpha+t\bar{\lambda})\bar{v}$, and $(\mathcal{H}+t\hat{\mathcal{K}}_\mu)x = \mathcal{H}x$ for any $x \in \ker(\mathcal{H} - \lambda I)^{2n}$ and $\lambda \in \Lambda(\mathcal{H}) \setminus \{i\alpha, -i\alpha\}$.

Proof. Since the Hermitian form $(x, y) \mapsto -ix^\star Jy$ is non-degenerate on $\ker(\mathcal{H} - i\alpha I)^{2n}$, there exists $w \in \ker(\mathcal{H} - i\alpha I)^{2n}$ such that $w^\star Jv = 1$. Hence $\hat{\mathcal{E}}_\mu$ is well defined. Obviously, $\hat{\mathcal{E}}_\mu v = \mu v$, whence $(\mathcal{H} + t\mathcal{E}_\mu)v = (i\alpha + t\mu)v$. Since $\ker(\mathcal{H} - i\alpha I)^{2n}$ is J-orthogonal to the other generalized eigenspaces of $\mathcal{H}$, we have $v^\star Jx = w^\star Jx = 0$ for any $x \in \ker(\mathcal{H} - \lambda I)^{2n}$ and $\lambda \in \Lambda(\mathcal{H}) \setminus \{i\alpha\}$. Thus $\hat{\mathcal{E}}_\mu x = 0$. This completes the proof of i). Assertion ii) is obvious, and iii) follows from the identity $\ker(\mathcal{H} + i\alpha I)^{2n} = \overline{\ker(\mathcal{H} - i\alpha I)^{2n}}$ and the J-orthogonality of the generalized eigenspaces. □

For the construction of Hamiltonian matrices that move eigenvalues off the imaginary axis, we need a J-neutral eigenvector. We now address the issue of existence of J-neutral eigenvectors corresponding to an imaginary eigenvalue of a Hamiltonian matrix. First, we show that a J-neutral eigenvector of $\mathcal{H}$ corresponding to an imaginary eigenvalue exists if the eigenvalue is defective.

Proposition 4.7 *Suppose that $v_1, v_2 \ldots, v_\ell$, $\ell \geq 2$, is a Jordan chain of the Hamiltionian matrix $\mathcal{H}$ associated with an imaginary eigenvalue $i\alpha$, i.e., $\mathcal{H}v_k = i\alpha\, v_k + v_{k-1}$ for $k = 1, \ldots, \ell$, where $v_0 := 0$. Then the subspace* $\mathrm{span}\{v_1, \ldots, v_{\lfloor \ell/2 \rfloor}\}$ *is J-neutral. In particular the eigenvector v_1 is J-neutral.*

Proof. We have $J(\mathcal{H} - i\alpha I) = -(\mathcal{H} - i\alpha I)^\star J$, $v_k = (\mathcal{H} - i\alpha I)^{\ell-k} v_\ell$ for $k = 1, \ldots, \ell$, and $(\mathcal{H} - i\alpha I)^q v_\ell = 0$ for $q \geq \ell$. Hence, if $k + j \leq \ell$, then $v_j^\star J v_k = v_\ell^\star ((\mathcal{H} - i\alpha I)^\star)^{\ell-j} J(\mathcal{H} - i\alpha I)^{\ell-k} v_\ell = (-1)^{\ell-j} v_\ell^\star J(\mathcal{H} - i\alpha I)^{2\ell-k-j} v_\ell = 0$. □

Proposition 4.7 shows that the first vector in a Jordan chain of length at least 2 is a J-neutral vector, but this may or may not be true for semi-simple purely imaginary eigenvalues. To characterize when this is the case, we need the sign characteristic of the purely imaginary eigenvalue (see Theorem 2.3), which allows to classify the purely imaginary eigenvalues into three distinct groups.

Definition 4.8 *Let $i\alpha$ be an imaginary eigenvalue of $\mathcal{H}$. Let X be a full column rank matrix such that* $\mathrm{span}(X) = \ker((\mathcal{H} - i\alpha I)^{2n})$. *Consider the matrix $Z := -iX^\star JX$. Then $i\alpha$ is said to have* positive sign characteristic, negative sign characteristic, or mixed sign characteristics, *depending on whether Z is positive definite, negative definite or indefinite, respectively.*

The following result characterizes the existence of a J-neutral eigenvector of a Hamiltonian matrix corresponding to an imaginary eigenvalue.

Proposition 4.9 *Let $i\alpha$ be an imaginary eigenvalue of a Hamiltonian matrix $\mathcal{H}$. Then $\mathcal{H}$ has a J-neutral eigenvector corresponding to $i\alpha$ if and only if $i\alpha$ has mixed sign characteristics.*

Proof. Recall that the Hermitian form $(x, y) \mapsto -ix^\star Jy$ is non-degenerate on $\ker(\mathcal{H} - i\alpha I)^{2n}$ and hence the matrix $Z = -iX^\star JX$ in Definition 4.8 is nonsingular. Suppose that there exists a J-neutral eigenvector associated with $i\alpha$. Then clearly Z is indefinite. Hence $i\alpha$ has mixed sign characteristics.

Conversely, suppose that $i\alpha$ has mixed sign characteristics, that is, Z is indefinite. By Proposition 4.7, a J-neutral eigenvector exists if the eigenvalue $i\alpha$ is defective. So, suppose that $i\alpha$ is semi-simple. Since Z is indefinite, there exist eigenvectors v_0 and v_1 such that $-iv_0^\star Jv_0 > 0$ and $-iv_1^\star Jv_1 < 0$. Hence by continuity there exists an eigenvector v of the form $v = \cos(t)v_0 + \sin(t)v_1$, for some $t \in \mathbb{R}$, such that $v^\star Jv = 0$. □

Note that if an imaginary eigenvalue of a Hamiltonian matrix is simple then it has either positive or negative sign characteristic. Indeed, when $i\alpha$ is simple, we have $Z = -ix^\star Jx$ which is either positive or negative. In fact, assuming that $x^\star x = 1$ we have $Z = s_j$, where $s_j = \pm 1$ is given by Theorem 2.3. Hence if $i\alpha$ has mixed sign characteristics, then $i\alpha$ is necessarily multiple. Note, further, that if $i\alpha$ is defective then by Proposition 4.7, $i\alpha$ has mixed sign characteristics. However, when $i\alpha$ is a non-defective multiple eigenvalue, it may or may not have mixed sign characteristics, see Example 3.1.

5 Minimal Hamiltonian perturbations

In this section we investigate how to move purely imaginary eigenvalues which are neither defective nor have mixed sign characteristics off the imaginary axis by suitable Hamiltonian perturbations. We begin with the problem of moving an eigenvalue of a Hamiltonian matrix to a specified point in the complex plane by a minimal Hamiltonian perturbation. This will play an important role in moving eigenvalues to specific points outside a strip S_τ as required in Problem **B**.

We have already seen that a purely imaginary eigenvalue can be moved off the imaginary axis by a small Hamiltonian perturbation if and only if it has a J-neutral eigenvector and that such an eigenvector exists if and only if the eigenvalue has mixed sign characteristic. Also, we have seen that a purely imaginary eigenvalue has mixed sign characteristics if and only if the eigenvalue is associated with a J-neutral eigenvector.

In order to move a purely imaginary eigenvalue having either positive or negative sign characteristic from the imaginary axis by a Hamiltonian perturbation, we therefore first need to coalesce it with another purely imaginary eigenvalue of opposite sign characteristic.

Thus, in this case we split the construction of perturbations that move the eigenvalues off the imaginary axis into two steps. First, we construct a minimal Hamiltonian perturbation that coalesces two eigenvalues having negative and positive sign characteristics into an imaginary eigenvalue having mixed sign characteristics. This moves the eigenvalues on the boundary of the set required in Problem **A**. Second, we move the resulting imaginary eigenvalue with mixed sign characteristics off the imaginary axis by a small Hamiltonian perturbation as required in Problem **B**.

Since we have already addressed the second stage of the problem in the previous section, we now address the first step of the construction.

For this purpose, we make use of both the backward error for the Hamiltonian eigenvalue problem and of Hamiltonian pseudospectra. These quantities are introduced and discussed in the following subsections. In the third subsection we then determine perturbations of minimum norm which remove a pair of eigenvalues from the imaginary axis.

5.1 Backward errors

We begin with the construction of backward errors for eigenvalues of a Hamiltonian matrix. The *Hamiltonian backward error* associated with a complex number $\lambda \in \mathbb{C}$ is defined by

$$\eta^{\mathrm{Ham}}(\lambda, \mathcal{H}) := \inf\{\, \|\mathcal{E}\| : \ \mathcal{E} \in \mathbb{F}^{2n,2n} \text{ Hamiltonian}, \ \lambda \in \Lambda(\mathcal{H} + \mathcal{E})\}. \tag{12}$$

Note that in general $\eta(z, \mathcal{H})$ will be different for $\mathbb{F} = \mathbb{C}$ and for $\mathbb{F} = \mathbb{R}$. We use the notation $\eta_F^{\mathrm{Ham}}(\lambda, \mathcal{H})$ and $\eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H})$, when the norm in (12) is the Frobenius norm and the spectral norm, respectively.

Theorem 5.1 *Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be a Hamiltonian matrix, and let $\lambda \in \mathbb{C}$ be such that $\mathrm{Re}\,\lambda \neq 0$. Then we have*

$$\begin{aligned} \eta_F^{\mathrm{Ham}}(\lambda, \mathcal{H}) &= \min_{\|x\|_2=1} \Big\{ \sqrt{2\|(\mathcal{H} - \lambda I)x\|_2^2 - |x^\star J \mathcal{H} x|^2} \\ &\qquad : x \in \mathbb{C}^{2n}, x^\star J x = 0 \Big\}, \qquad (13) \\ \eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H}) &= \min_{\|x\|_2=1} \{\|(\mathcal{H} - \lambda I)x\|_2 : x \in \mathbb{C}^{2n}, \ x^\star J x = 0\}. \qquad (14) \end{aligned}$$

In particular, we have $\eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H}) \leq \eta_F^{\mathrm{Ham}}(\lambda, \mathcal{H}) \leq \sqrt{2}\,\eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H})$.

Suppose that the minima in (13), and (14) are attained for $u \in \mathbb{C}^{2n}$ and $v \in \mathbb{C}^{2n}$, respectively. Let $\mathcal{E} := \mathcal{G}(u, (\lambda I - \mathcal{H})u)$ and $\mathcal{K} := \mathcal{F}(v, (\lambda I - \mathcal{H})v)$, where $\mathcal{G}$ and $\mathcal{F}$ are as in Theorem 3.3. Then

$$\begin{aligned} \|\mathcal{E}\|_F &= \eta_F^{\mathrm{Ham}}(\lambda, \mathcal{H}) \text{ and } (\mathcal{H} + \mathcal{E})u = \lambda u, \\ \|\mathcal{K}\|_2 &= \eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H}) \text{ and } (\mathcal{H} + \mathcal{K})v = \lambda v. \end{aligned}$$

Proof. Let $x \in \mathbb{C}^n$ be nonzero. Then by Theorem 3.3 there exists a Hamiltonian matrix $\mathcal{E} \in \mathbb{C}^{2n,2n}$ such that $(\mathcal{H} + \mathcal{E})x = \lambda x$ if and only if $x^\star J x = 0$. Indeed, setting $r = \lambda x - \mathcal{H}x$, it follows that $x^\star J r$ is real if and only if $x^\star J x = 0$. So, suppose that $x^\star J x = 0$ and w.l.o.g. that $x^\star x = 1$. Then by Theorem 3.3, $\mathcal{E} := \mathcal{G}(x, r)$ is the unique Hamiltonian matrix such that $(\mathcal{H} + \mathcal{E})x = \lambda x$ and $\mathcal{E}$ has minimal Frobenius norm given by

$$\|\mathcal{E}\|_F = \sqrt{2\|(\mathcal{H} - \lambda I)x\|_2^2 - |x^\star J(\mathcal{H} - \lambda I)x|^2}.$$

Similarly, by Theorem 3.3, $\mathcal{K} := \mathcal{F}(x, r)$ is a Hamiltonian matrix such that $(\mathcal{H} + \mathcal{K})x = \lambda x$ and $\mathcal{K}$ has minimal spectral norm given by

$$\|\mathcal{K}\|_2 = \|(\mathcal{H} - \lambda I)x\|_2.$$

Then the claim follows by taking the minimum over all $x \in \mathbb{C}^{2n}$ such that $x^\star J x = 0$. $\square$

Note that it is a nontrivial task to determine the minimal values $\eta_2^{\mathrm{Ham}}(\lambda, \mathcal{H})$ and $\eta_F^{\mathrm{Ham}}(\lambda, \mathcal{H})$, when $\lambda \in \mathbb{C}$ and $\mathrm{Re}\,\lambda \neq 0$. In contrast, it is relatively simple to determine these minimal values for purely imaginary values $\lambda = i\omega$ with $\omega \in \mathbb{R}$. The construction in Proposition 5.3 below is based on the following observation.

Observation 5.2 Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be Hamiltonian, and let $\lambda_1, \ldots, \lambda_{2n} \in \mathbb{R}$ denote the eigenvalues of the Hermitian matrix $J\mathcal{H}$. Let $v_1, \ldots, v_{2n} \in \mathbb{C}^{2n}$ be an orthonormal basis of eigenvectors of $J\mathcal{H}$, such that $J\mathcal{H}\, v_k = \lambda_k\, v_k$. Then $|\lambda_1|, \ldots, |\lambda_{2n}|$ are the singular values of $\mathcal{H}$ and the vectors v_k are the associated right singular vectors. The associated left singular vectors are $u_k = -\mathrm{sign}(\lambda_k)\, Jv_k$. Indeed, the matrices $V = [v_1, \ldots, v_{2n}]$, $U = [u_1, \ldots, u_{2n}]$ are unitary, and from $J\mathcal{H}\, V = V\, \mathrm{diag}(\lambda_1, \ldots, \lambda_{2n})$ it follows that $\mathcal{H} = U\mathrm{diag}(|\lambda_1|, \ldots, |\lambda_{2n}|)\, V^\star$.

Proposition 5.3 *Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be Hamiltonian and $\omega \in \mathbb{R}$. Let v be a normalized eigenvector of the Hermitian matrix $J(\mathcal{H} - i\omega I)$ corresponding to an eigenvalue $\lambda \in \mathbb{R}$. Then $|\lambda|$ is a singular value of the Hamiltonian matrix $\mathcal{H} - i\omega I$ and v is an associated right singular vector.*

Further, the matrices

$$\begin{aligned} \mathcal{E} &= \lambda J v v^\star, && (15)\\ \mathcal{K} &= \lambda J [v\ \bar{v}]\, [v\ \bar{v}]^+ && (16) \end{aligned}$$

are Hamiltonian, $\mathcal{K}$ is real and we have $(\mathcal{H} + \mathcal{E})v = (\mathcal{H} + \mathcal{K})v = i\omega v$. Furthermore, $\|\mathcal{E}\|_2 = \|\mathcal{K}\|_2 = |\lambda|$ and $\|\mathcal{K}\|_F = \sqrt{2}\, \|\mathcal{E}\|_F = \sqrt{2}\, |\lambda|$.

Moreover, suppose that λ is an eigenvalue of $J(\mathcal{H} - i\omega I)$ of smallest absolute value and let $\sigma_{\min}(\mathcal{H} - i\omega I)$ be the smallest singular value of $\mathcal{H} - i\omega I$. Then $|\lambda| = \sigma_{\min}(\mathcal{H} - i\omega I)$ and we have

$$\begin{aligned} \eta_F^{\mathrm{Ham}}(i\omega, \mathcal{H}) &= \eta_2^{\mathrm{Ham}}(i\omega, \mathcal{H}) = |\lambda| = \|\mathcal{E}\|_2, \ \textit{when } \mathbb{F} = \mathbb{C},\\ \eta_F^{\mathrm{Ham}}(i\omega, \mathcal{H}) &= \sqrt{2}\, \eta_2^{\mathrm{Ham}}(i\omega, \mathcal{H}) = \sqrt{2}|\lambda| = \|\mathcal{K}\|_F, \ \textit{when } \mathbb{F} = \mathbb{R}. \end{aligned}$$

Proof. The first assertion follows by applying Observation 5.2 to the Hamiltonian matrix $\mathcal{H} - i\omega I$. By construction, $\mathcal{H}$ and $\mathcal{K}$ are Hamiltonian and $(\mathcal{H} + \mathcal{E})v = (\mathcal{H} + \mathcal{K})v = i\omega v$. Note that by Lemma 3.4, $\mathcal{K}$ is real. Obviously, we have $\|\mathcal{E}\|_2 = \|\mathcal{K}\|_2 = |\lambda|$ and $\|\mathcal{K}\|_F = \sqrt{2}\|\mathcal{E}\|_F = \sqrt{2}|\lambda|$.

Now if λ has the smallest absolute value then by Observation 5.2 we have $\sigma_{\min}(\mathcal{H} - i\omega I) = |\lambda|$. Since $\|\mathcal{K}\|_2 = \|\mathcal{E}\|_F = \|\mathcal{E}\|_2 = \sigma_{\min}(\mathcal{H} - i\omega I)$ and $\eta_F^{\mathrm{Ham}}(i\omega, \mathcal{H}) \geq \eta_2^{\mathrm{Ham}}(i\omega, \mathcal{H}) \geq \sigma_{\min}(\mathcal{H} - i\omega I)$, the desired result follows for the case when $\mathbb{F} = \mathbb{C}$.

For the Frobenius norm and the case when $\mathbb{F} = \mathbb{R}$, by Corollary 3.5 we conclude that $\mathcal{K} = \mathcal{G}([v, \bar{v}],\, \lambda J[v, \bar{v}])$ and that $\|\mathcal{K}\|_F = \sqrt{2}\|\mathcal{K}\|_2 = \sqrt{2}|\lambda| = \sqrt{2}\, \sigma_{\min}(\mathcal{H} - i\omega I)$. Hence the desired result follows. □

Proposition 5.3 in particular states that a Hamiltonian perturbation of $\mathcal{H}$ of smallest norm that moves an eigenvalue to the point $i\omega$ can be constructed from an eigenpair (v, λ) of $J(\mathcal{H} - i\omega I)$, where λ has the smallest absolute value. Our next results shows that the eigenpair (v, λ) can be chosen as a piecewise analytic (but not necessarily continuous) function of ω.

Proposition 5.4 *Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be Hamiltonian, let $F(\omega) = J(\mathcal{H} - i\omega I)$ and $f(\omega) = \sigma_{\min}(\mathcal{H} - i\omega I)$ for $\omega \in \mathbb{R}$. There exist a finite number ℓ of real values $\gamma_1 < \gamma_2 < \ldots < \gamma_\ell$ and functions $\lambda_{\min} : \mathbb{R} \to \mathbb{R}$, $v : \mathbb{R} \to \mathbb{C}^{2n}$ which are analytic on $\mathbb{R} \setminus \{\gamma_1, \ldots, \gamma_\ell\}$ and have the following properties.*

a) $F(\omega)v(\omega) = \lambda_{\min}(\omega)v(\omega)$, $|\lambda_{\min}(\omega)| = \min\{\, |\lambda| : \lambda \in \Lambda(F(\omega))\}$ and $\|v(\omega)\|_2 = 1$ for all $\omega \in \mathbb{R}$.

b) For each $k \in \{0, 1, \ldots, \ell\}$ either $\lambda_{\min}(\omega) = f(\omega)$ for all $\omega \in (\gamma_k, \gamma_{k+1})$ or $\lambda_{\min}(\omega) = -f(\omega)$ for all $\omega \in (\gamma_k, \gamma_{k+1})$, where we set $\gamma_0 = -\infty$ and $\gamma_{\ell+1} = \infty$.

c) The vector $v(\omega)$ is a right singular vector of the matrix $\mathcal{H} - i\omega I$ associated with the smallest singular value.

d) The derivative of $\lambda_{\min}(\cdot)$ at $\omega \in \mathbb{R} \setminus \{\gamma_1, \ldots, \gamma_\ell\}$ satisfies

$$\lambda'_{\min}(\omega) \quad = \quad -iv(\omega)^\star Jv(\omega).$$

e) At each of the (exceptional) points γ_k the left and the right limits of $\lambda_{\min}(\cdot)$ and $v(\cdot)$ exist. Suppose that $\lambda_{\min}(\cdot)$ is continuous at γ_k. Then the left and the right side derivative of $\lambda_{\min}(\cdot)$ at γ_k both exist and satisfy

$$\lim_{\omega \to \gamma_k \pm} \frac{\lambda_{\min}(\omega) - \lambda_{\min}(\gamma_k)}{\omega - \gamma_k} = \lim_{\omega \to \gamma_k \pm} \lambda'_{\min}(\omega).$$

Proof. Note that $F(\omega) = J(\mathcal{H} - i\omega I)$, $\omega \in \mathbb{R}$, is a Hermitian matrix. By [37, pp. 29-33] there exist analytic functions $\omega \mapsto v_1(\omega), \ldots, v_{2n}(\omega) \in \mathbb{C}^{2n}$ and $\omega \mapsto \lambda_1(\omega), \ldots, \lambda_{2n}(\omega) \in \mathbb{R}$ such that for each ω the vectors $v_j(\omega)$ form an orthonormal basis of $\mathbb{C}^{2n}$ and $F(\omega)v_j(\omega) = \lambda_j(\omega)v_j(\omega)$. The derivative of λ_j at ω satisfies

$$\begin{aligned}
\lambda'_j(\omega) &= \frac{d}{d\omega}\left(v_j(\omega)^\star F(\omega) v_j(\omega)\right) \\
&= v_j(\omega)^\star F'(\omega) v_j(\omega) + v'_j(\omega)^\star F(\omega) v_j(\omega) + v_j(\omega)^\star F(\omega) v'_j(\omega) \\
&= -iv_j(\omega)^\star J v_j(\omega) + \lambda_j(\omega) \underbrace{\left(v'_j(\omega)^\star v_j(\omega) + v_j(\omega)^\star v'_j(\omega)\right)}_{=\frac{d}{d\omega}\|v_j(\omega)\|^2 = 0} \\
&= -iv_j(\omega)^\star J v_j(\omega). \qquad (17)
\end{aligned}$$

For each pair of indices j, k the analytic functions $\lambda_j(\cdot), \lambda_k(\cdot)$ are either identical or meet in a discrete set $P_{j,k} \subset \mathbb{R}$. Analogously the functions $-\lambda_j(\cdot), \lambda_k(\cdot)$, are either identical or meet in a discrete set $Q_{j,k} \subset \mathbb{R}$. Since the union of the graphs of the functions $\pm\lambda_j(\cdot)$ equals the algebraic curve $\{(\omega, \lambda) \in \mathbb{R}^2 \mid \det((F(\omega) - \lambda I)(F(\omega) + \lambda I)) = 0\}$, both the sets $P_{j,k}$ and $Q_{j,k}$ are finite. Let $\{\gamma_1, \ldots, \gamma_r\}$, $\gamma_k < \gamma_{k+1}$, denote the union of the sets P_{jk} and the sets Q_{jk}. By the third claim of Proposition 5.3, we have that $f(\omega) = \min_{j=1,\ldots,2n} |\lambda_j(\omega)|$. It follows that to each interval $\mathcal{I}_k = (\gamma_k, \gamma_{k+1})$ there exists an index j such that either $\lambda_j(\omega) = f(\omega)$ for all $\omega \in \mathcal{I}_k$ or $\lambda_j(\omega) = -f(\omega)$ for all $\omega \in \mathcal{I}_k$. Define $\lambda_{\min}(\omega) := \lambda_j(\omega)$, $v(\omega) := v_j(\omega)$ for $\omega \in \mathcal{I}_k$ and $\lambda_{\min}(\gamma_k) := \lambda_j(\gamma_k)$, $v(\gamma_k) := v_j(\gamma_k)$. Then the functions $\lambda_{\min}(\cdot)$ and $v(\cdot)$ have the required properties. $\square$

Example 5.5 The upper diagram of Figure 1 shows the eigenvalue curves $\omega \mapsto \lambda_j(\omega)$ of the Hermitian matrix function $\omega \mapsto J(\mathcal{H}_1 - i\omega I)$ for $\omega \in [-16, 16]$ and $\mathcal{H}_1 := \begin{bmatrix} 0 & G_1 \\ H_1 & 0 \end{bmatrix}$, where

$$G_1 := \begin{bmatrix} 7 & -4 & 2 & -11 & 0 \\ -4 & -37 & 31 & -8 & 0 \\ 2 & 31 & -28 & 4 & 0 \\ -11 & -8 & 4 & 28 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{bmatrix}, \quad H_1 := \begin{bmatrix} 11 & 16 & 16 & 5 & 0 \\ 16 & 21 & 30 & 8 & 0 \\ 16 & 30 & 48 & 8 & 0 \\ 5 & 8 & 8 & -1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}.$$

The spectrum of $\mathcal{H}_1$ is $\Lambda(\mathcal{H}_1) = \{\pm 3i, \pm 5i \pm 10i, \pm 15i\}$ and the eigenvalues $\pm 10i$ have multiplicity 2, while the other eigenvalues are simple. At the real parts of the eigenvalues of $\mathcal{H}_1$ the eigenvalue curves $\lambda_j(\cdot)$ cross the real axis. Observe that, according to (17), the sign characteristics of the eigenvalues of $\mathcal{H}_1$ can be read off from the slopes of the curves $\lambda_j(\cdot)$. The λ_j-curves crossing the real axis at $-15, -3$ and 5 have positive slope, i.e., the eigenvalues $-15i, -3i$ and $5i$ have positive sign characteristic. The $\lambda_j(\cdot)$-curves crossing the real axis at $-5, 3$ and 15 have negative slope i.e., the eigenvalues $-5i, 3i$ and $15i$ have negative sign characteristic. At the points ± 10 there are two λ_j-curves crossing the real axis with positive and negative slopes. Thus, the eigenvalues $\pm 10i$ both have mixed sign characteristic. The graph of the function $\omega \mapsto \lambda_{\min}(\omega) = \lambda_{\min}(J(\mathcal{H}_1 - i\omega I))$ from Proposition 5.4 is depicted by thick curves. Note that this function is piecewise analytic but discontinuous. The lower diagram of Figure 1 shows the singular value curves of the pencil $\omega \mapsto \mathcal{H}_1 - i\omega I$. The graph of the continuous function $\omega \mapsto \sigma_{\min}(\mathcal{H}_1 - i\omega I)$ is depicted as a thick curve. Note that $\sigma_{\min}(\mathcal{H}_1 - i\omega I) = |\lambda_{\min}(\omega)|$.

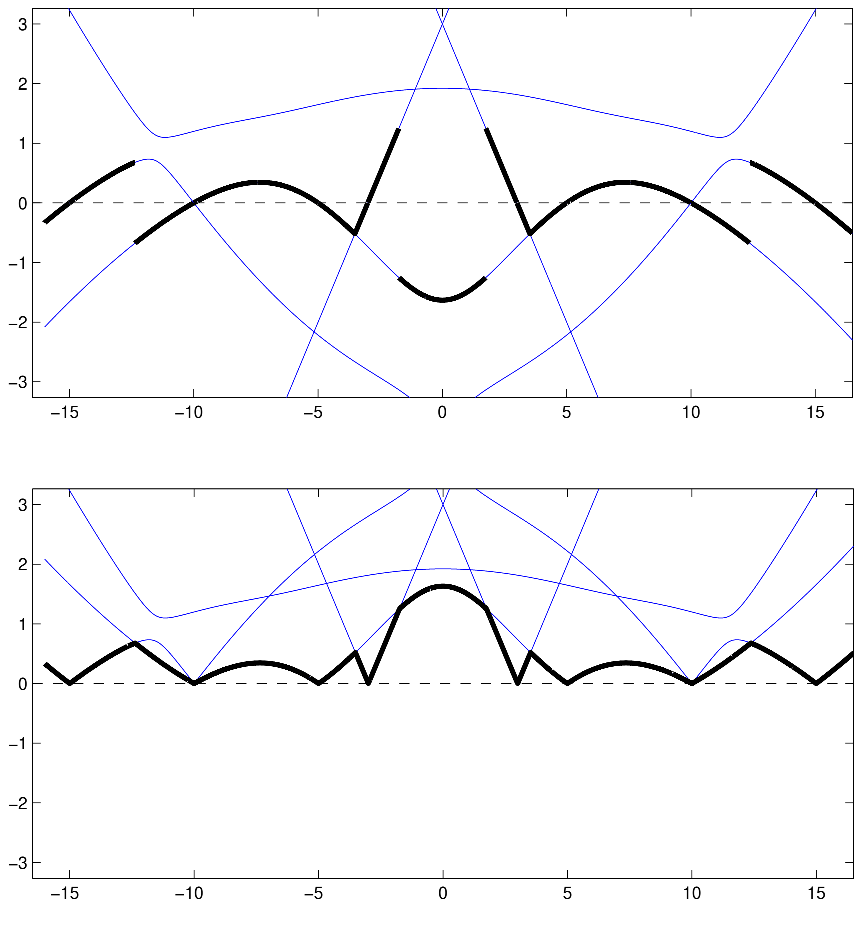

Figure 1: Eigenvalue and singular value curves for Example 5.5

The following proposition characterizes the existence of J-neutral eigenvectors in terms of the local extrema of the eigenvalue curves.

Proposition 5.6 *Suppose the function $\lambda_{\min} : \mathbb{R} \to \mathbb{R}$ of Proposition 5.4 is continuous at $\omega_0 \in \mathbb{R}$ and attains a local extremum at ω_0. Then there exists a J-neutral normalized eigenvector v_0 of the Hermitian matrix $J(\mathcal{H} - i\omega_0 I)$ corresponding to the eigenvalue $\lambda_{\min}(\omega_0)$.*

Proof. If $\omega_0 \in \mathbb{R} \setminus \{\gamma_1, \dots, \gamma_r\}$ then the derivative of $\lambda_{\min}(\cdot)$ at ω_0 satisfies $0 = \lambda'_{\min}(\omega_0) = -iv(\omega_0)^\star Jv(\omega_0)$. Hence, $v_0 := v(\omega_0)$ is J-neutral if $\lambda_{\min}$ attains a local extremum at ω_0. Suppose now that $\omega_0 \in \{\gamma_1, \dots \gamma_r\}$. Assume w.l.o.g. that ω_0 is a local maximum. Then the left sided derivative of $\lambda_{\min}(\cdot)$ at ω_0 is nonnegative and the right sided derivative is non-positive. Hence, it follows from claim e) of Proposition 5.4 that

$$\begin{array}{rcl} 0 & \le & \lim\limits_{\omega \to \omega_0 -} \lambda'_{\min}(\omega) = \lim\limits_{\omega \to \omega_0 -} (-iv(\omega)^\star Jv(\omega)) = -iv_-^\star Jv_-, \\ 0 & \ge & \lim\limits_{\omega \to \omega_0 +} \lambda'(\omega) = \lim\limits_{\omega \to \omega_0 -} (-iv(\omega)^\star Jv(\omega)) = -iv_+^\star Jv_+, \end{array}$$

where $v_\pm = \lim_{\omega \to \omega_0 \pm} v(\omega)$. Suppose that v_+ and v_- are linearly dependent. Then $-iv_-^\star Jv_- = -iv_+^\star Jv_+ = 0$, i.e., $v_0 := v_+$ has the required properties. If v_+ and v_- are linearly independent, then let $u_t = tv_+ + (1-t)v_-$. In this case for all $t \in \mathbb{R}$, $u_t \neq 0$ and $J(\mathcal{H} - i\omega_0 I)u_t = \lambda_{\min}(\omega_0)u_t$. Furthermore, $-iu_0^\star Ju_0 \le 0$ and $-iu_1^\star Ju_1 \ge 0$. By continuity there exists $t_0 \in [0,1]$ such that $-iu_{t_0}^\star Ju_{t_0} = 0$, and hence, $v_0 := u_{t_0}/\|u_{t_0}\|$ has the required properties. □

5.2 Pseudospectra

Let $A \in \mathbb{C}^{n,n}$ and let $\epsilon \ge 0$. Then the *ϵ-pseudospectrum* of A is defined as

$$\Lambda_\epsilon(A, \mathbb{F}) = \bigcup_{\|E\|_2 \le \epsilon} \{\Lambda(A + E) : E \in \mathbb{F}^{n,n}\}.$$

It is well-known [43] that in the complex case when $\mathbb{F} = \mathbb{C}$, we have

$$\Lambda_\epsilon(A, \mathbb{C}) = \{ z \in \mathbb{C} : \ \sigma_{\min}(A - z I) \le \epsilon \},$$

where $\sigma_{\min}(\cdot)$ denotes the minimum singular value. Since we are interested in structured perturbations, we also consider the *Hamiltonian ϵ-pseudospectrum* defined by

$$\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) = \bigcup_{\|\mathcal{E}\|_2 \le \epsilon} \{\Lambda(\mathcal{H} + \mathcal{E}) : \mathcal{E} \in \mathbb{F}^{2n,2n} \text{ and } (J\mathcal{E})^\star = J\mathcal{E} \}.$$

It is obvious that

$$\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{C}) = \{ z \in \mathbb{C} : \ \eta_2^{\mathrm{Ham}}(z, \mathcal{H}) \le \epsilon \},$$

where $\eta_2^{\mathrm{Ham}}(z, \mathcal{H})$ is the Hamiltonian backward error as defined in (12).

Note that the pseudospectra so defined will in general be different for $\mathbb{F} = \mathbb{C}$ and for $\mathbb{F} = \mathbb{R}$, however, for purely imaginary eigenvalues, the following result is an immediate consequence of Proposition 5.3.

Corollary 5.7 *Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be Hamiltonian. Consider the pseudospectra $\Lambda_\epsilon(\mathcal{H}; \mathbb{F})$ and $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}; \mathbb{F})$. Then,*

$$\begin{array}{rcl} \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}; \mathbb{C}) \cap i\mathbb{R} & = & \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}; \mathbb{R}) \cap i\mathbb{R} = \Lambda_\epsilon(\mathcal{H}; \mathbb{C}) \cap i\mathbb{R} = \Lambda_\epsilon(\mathcal{H}; \mathbb{R}) \cap i\mathbb{R} \\ & = & \{i\omega \ : \ \omega \in \mathbb{R}, \ \sigma_{\min}(\mathcal{H} - i\omega I) \le \epsilon\} \\ & = & \{i\omega \ : \ \omega \in \mathbb{R}, \ |\lambda_{\min}(J(\mathcal{H} - i\omega I))| \le \epsilon\}, \end{array}$$

where $\lambda_{\min}(\cdot)$ denotes the eigenvalue function from Proposition 5.4.

In Definition 4.8 we have associated sign characteristics to the purely imaginary eigenvalues of a Hamiltonian matrix. We now associate sign characteristics to the connected components of a Hamiltonian pseudospectrum.

Definition 5.8 *Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$. Then a connected component $\mathcal{C}_\epsilon(\mathcal{H})$ of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ is said to have positive (resp., negative) sign characteristic if for all Hamiltonian perturbations $\mathcal{E}$ with $\|\mathcal{E}\|_2 \leq \epsilon$ each eigenvalue of $\mathcal{H} + \mathcal{E}$ that is contained in $\mathcal{C}_\epsilon(\mathcal{H})$ has positive (resp., negative) sign characteristic.*

Observe that if a component $\mathcal{C}_\epsilon(\mathcal{H})$ of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ has positive (resp., negative) sign characteristic then $\mathcal{C}_\epsilon(\mathcal{H}) \subset i\mathbb{R}$ and all eigenvalues of $\mathcal{H}$ that are contained in $\mathcal{C}_\epsilon(\mathcal{H})$ have positive (resp., negative) sign characteristic. We now show that the sign characteristic of $\mathcal{C}_\epsilon(\mathcal{H})$ is completely determined by the sign characteristic of the eigenvalues of $\mathcal{H}$ that are contained in $\mathcal{C}_\epsilon(\mathcal{H})$.

Theorem 5.9 *Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ and $\mathcal{C}_\epsilon(\mathcal{H})$ be a connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$. For a Hamiltonian matrix $\mathcal{E} \in \mathbb{F}^{2n,2n}$ with $\|\mathcal{E}\|_2 \leq \epsilon$, let $X_{\mathcal{E}}$ be a full column rank matrix whose columns form a basis of the direct sum of the generalized eigenspaces $\ker(\mathcal{H} + \mathcal{E} - \lambda I)^{2n}$, $\lambda \in \mathcal{C}_\epsilon(\mathcal{H}) \cap \Lambda(\mathcal{H} + \mathcal{E})$. Set $Z_{\mathcal{E}} := -iX_{\mathcal{E}}^\star J X_{\mathcal{E}}$. Then the following conditions are equivalent.*

a) The component $\mathcal{C}_\epsilon(\mathcal{H})$ has positive (resp., negative) sign characteristic.

b) All eigenvalues of $\mathcal{H}$ that are contained in $\mathcal{C}_\epsilon(\mathcal{H})$ have positive (resp., negative) sign characteristic.

c) The matrix Z_0 associated with $\mathcal{E} = 0$ is positive (resp., negative) definite.

d) The matrix $Z_{\mathcal{E}}$ is positive (resp., negative) definite for all Hamiltonian matrix $\mathcal{E}$ with $\|\mathcal{E}\|_2 \leq \epsilon$.

Proof. Without loss of generality suppose that $\mathcal{C}_\epsilon(\mathcal{H})$ has positive sign characteristic. Then obviously all eigenvalues of $\mathcal{H}$ that are contained in $\mathcal{C}_\epsilon(\mathcal{H})$ have positive sign characteristic. This proves a) $\Rightarrow$ b).

Next, suppose that $\Lambda(\mathcal{H}) \cap \mathcal{C}_\epsilon(\mathcal{H})$ contains p distinct eigenvalues $i\alpha_1, \ldots, i\alpha_p$ each of which has positive sign characteristic. Let X_k be a full column rank matrix whose columns form a basis of $\ker(\mathcal{H} - i\alpha_k)^{2n}$ for $k = 1, \ldots, p$. Then the columns of JX_k forms a basis of the left generalized eigenspace of $\mathcal{H}$ corresponding to the eigenvalue $i\alpha_k$. Hence $X_k^\star J X_l = -(JX_k)^\star X_l = 0$ for $l \neq k$. Since $i\alpha_k$ has positive sign characteristic, the matrix $-iX_k^\star J X_k$ is positive definite for $k = 1, \ldots, p$. Now considering $X := [X_1, \ldots, X_p]$ it follows that $-iX^\star JX = \mathrm{diag}(-iX_1^\star JX_1, \ldots, -iX_p^\star JX_p)$ is positive definite. Since $X_0 = XM$ for some nonsingular matrix M, it follows that Z_0 is congruent to $-iX^\star JX$. Hence Z_0 is positive definite. This proves b) $\Rightarrow$ c).

Now suppose that Z_0 is positive definite. Since $\mathcal{C}_\epsilon(\mathcal{H})$ is a closed and connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$, there is a simple closed rectifiable curve Γ such that $\Gamma \cap \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) = \emptyset$ and that the component $\mathcal{C}_\epsilon(\mathcal{H})$ lies inside the curve Γ. Let $\mathcal{E}$ be a Hamiltonian matrix with $\|\mathcal{E}\|_2 \leq \epsilon$. Consider the matrix $\mathcal{H}(t) := \mathcal{H} + t\mathcal{E}$ for $t \in \mathbb{C}$. Then by [II.3-II.4, page 66-68, [20]] there exists a matrix $X_{\mathcal{E}}(t)$ such that $X_{\mathcal{E}}(t)$ is analytic in $D_\Gamma := \{t \in \mathbb{C} : |t|\|\mathcal{E}\|_2 < \min_{z\in\Gamma} \sigma_{\min}(\mathcal{H} - zI)\}$. Further, for each $t \in D_\Gamma$, the matrix $X_{\mathcal{E}}(t)$ has full column rank and the columns form a basis of the direct sum of the generalized eigenspaces $\ker(\mathcal{H}(t) - \lambda I)^{2n}, \lambda \in \Lambda(\mathcal{H}(t)) \cap \mathcal{C}_\epsilon(\mathcal{H}) =: \sigma_{\mathcal{E}}(t)$. Since $\|\mathcal{E}\|_2 \leq \epsilon$ and $\min_{z\in\Gamma} \sigma_{\min}(\mathcal{H} - zI) > \epsilon$, it follows that

$[0,1] \subset D_\Gamma$. Hence the matrix $X_\mathcal{E}(t)$ is smooth on $[0,1]$. Set $Z_\mathcal{E}(t) := -iX_\mathcal{E}(t)^\star JX_\mathcal{E}(t)$ for $t \in [0,1]$. Then $Z_\mathcal{E}(t)$ is continuous and, by Corollary 2.2, $X_\mathcal{E}(t)$ is nonsingular for $t \in [0,1]$. Indeed, since $\sigma_\mathcal{E}(t)$ is symmetric with respect to the imaginary axis, the columns of $X_\mathcal{E}(t)$ form a basis of the direct sum of the J-nondegenerate and pairwise J-orthogonal subspaces $\ker(\mathcal{H}(t) - i\alpha\, I)^{2n}$, $i\alpha \in \sigma_\mathcal{E}(t)$, and $\ker(\mathcal{H}(t) - \lambda\, I)^{2n} \oplus \ker(\mathcal{H}(t) + \bar{\lambda}\, I)^{2n}$, $\lambda \in \sigma_\mathcal{E}(t) \setminus i\mathbb{R}$, see Corollary 2.2. It follows that $\mathrm{span}(X_\mathcal{E}(t))$ is J-nondegenerate. Thus, $Z_\mathcal{E}(t)$ is nonsingular for all $t \in [0,1]$. Since $Z_\mathcal{E}(0)$ is positive definite and $Z_\mathcal{E}(t)$ is nonsingular for all t in the connected set $[0,1]$, it follows that $Z_\mathcal{E}(t)$ is positive definite for all $t \in [0,1]$. This shows that $Z_\mathcal{E}$ is positive definite. Since $\mathcal{E}$ is arbitrary, we conclude that the assertion in d) holds. This proves c) $\Rightarrow$ d).

Finally, suppose that the assertion in d) holds. Then obviously for all Hamiltonian matrices $\mathcal{E}$ with $\|\mathcal{E}\|_2 \le \epsilon$, the eigenvalues in $\Lambda(\mathcal{H} + \mathcal{E}) \cap \mathcal{C}_\epsilon(\mathcal{H})$ are purely imaginary and have positive sign characteristic. In other words, $\mathcal{C}_\epsilon(\mathcal{H})$ has positive sign characteristic. This completes the proof. □

The following result is an immediate consequence of the proof of Theorem 5.9.

Corollary 5.10 *Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ and $\mathcal{C}_\epsilon(\mathcal{H})$ be a connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$. For a Hamiltonian matrix $\mathcal{E} \in \mathbb{F}^{2n,2n}$ with $\|\mathcal{E}\|_2 \le \epsilon$, let $X_\mathcal{E}$ be a full column rank matrix whose columns form a basis of the direct sum of the generalized eigenspaces $\ker(\mathcal{H} + \mathcal{E} - \lambda I)^{2n}$, $\lambda \in \mathcal{C}_\epsilon(\mathcal{H}) \cap \Lambda(\mathcal{H} + \mathcal{E})$. Set $Z_\mathcal{E} := -iX_\mathcal{E}^\star JX_\mathcal{E}$. Then the following holds.*

i) The rank of $\mathcal{X}_\mathcal{E}$ is constant for all Hamiltonian matrices $\mathcal{E}$ with $\|\mathcal{E}\|_2 \le \epsilon$.

ii) If $\mathcal{C}_\epsilon(\mathcal{E}) \cap i\mathbb{R} = \emptyset$ then $Z_\mathcal{E} = 0$ for all Hamiltonian matrices $\mathcal{E}$ with $\|\mathcal{E}\|_2 \le \epsilon$.

iii) If $\mathcal{C}_\epsilon(\mathcal{H}) \cap i\mathbb{R} \neq \emptyset$, then $\mathcal{C}_\epsilon(\mathcal{H}) = -\overline{\mathcal{C}_\epsilon(\mathcal{H})}$ and $Z_\mathcal{E}$ is nonsingular for all Hamiltonian matrices $\mathcal{E}$ with $\|\mathcal{E}\|_2 \le \epsilon$. Furthermore, the matrix $\mathcal{Z}_\mathcal{E}$ has the same inertia for all such $\mathcal{E}$.

iv) If $Z_\mathcal{E}$ is positive (resp., negative) definite for some Hamiltonian matrix $\mathcal{E}$ with $\|\mathcal{E}\|_2 \le \epsilon$ then $\mathcal{C}_\epsilon(\mathcal{H}) \subseteq i\mathbb{R}$ and $\mathcal{C}_\epsilon(\mathcal{H})$ has positive (resp., negative) sign characteristic.

The results in Theorem 5.9 and Corollary 5.10 provide an important insight into the evolution of purely imaginary eigenvalues of a Hamiltonian matrix subject to Hamiltonian perturbations. With a view to further understanding the evolution of purely imaginary eigenvalues of a Hamiltonian matrix, we now analyze the coalescence of pseudospectral components.

5.3 Coalescence of pseudospectral components

Consider the Hamiltonian pseudospectrum $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ of a Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$. Then obviously the set valued map $\epsilon \mapsto \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ is monotonically increasing, i.e., if $\epsilon < \delta$ then $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) \subset \Lambda_\delta^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$. Furthermore, for $\epsilon > 0$, the pseudospectrum $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ consists of at most $2n$ connected components and each component contains at least one eigenvalue of $\mathcal{H}$. Thus when ϵ is sufficiently small, then each component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ contains exactly one eigenvalue of $\mathcal{H}$ and as ϵ increases, these components expand in size and at some stage coalesce with each other. So, let $i\alpha$ be a purely imaginary eigenvalue of $\mathcal{H}$ and let $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ denote the connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ which contains $i\alpha$. Then for a sufficiently small ϵ the component $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ contains only $i\alpha$, i.e., $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \cap \Lambda(\mathcal{H}) = \{i\alpha\}$. Thus, if $i\alpha$ has either positive or negative sign characteristic, then by Theorem 5.9 we have

$\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \subset i\mathbb{R}$. This means that the eigenvalue $i\alpha$ cannot be removed from the imaginary axis by a Hamiltonian perturbation $\mathcal{E}$ of $\mathcal{H}$ such that $\|\mathcal{E}\|_2 \leq \epsilon$.

Next, let $i\beta$ be another purely imaginary eigenvalue of $\mathcal{H}$ with $\alpha < \beta$ and suppose that $\mathcal{C}_\epsilon(\mathcal{H}, i\beta)$ is a component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ containing $i\beta$ such that $\mathcal{C}_\epsilon(\mathcal{H}, i\beta) \cap \Lambda(\mathcal{H}) = \{i\beta\}$. Suppose further that $i\beta$ has either positive or negative sign characteristic so that by Theorem 5.9 we have $\mathcal{C}_\epsilon(\mathcal{H}, i\beta) \subset i\mathbb{R}$. Assume that $\mathcal{H}$ does not have an eigenvalue $i\gamma$ with $\gamma \in (\alpha, \beta)$ and that the component $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ coalesces with the component $\mathcal{C}_\epsilon(\mathcal{H}, i\beta)$ at $i\omega_0$ as ϵ tends to ϵ_0, i.e., $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \cap \mathcal{C}_\epsilon(\mathcal{H}, i\beta) = \emptyset$ for $\epsilon < \epsilon_0$ and $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \cap \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta) = \{i\omega_0\}$. We now investigate the geometry of the connected component $\mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\alpha) = \mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\beta)$ of $\Lambda_{\epsilon_0+\delta}^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ in a neighborhood of $i\omega_0$ for a small $\delta > 0$. In particular, we show that when $i\alpha$ and $i\beta$ have opposite sign characteristics, then the pseudospectrum $\Lambda_{\epsilon_0+\delta}^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ contains a disk centered at $i\omega_0$. Furthermore, in this case we show that there exists a Hamiltonian matrix $\mathcal{E}$ with $\|\mathcal{E}\|_2 = \epsilon_0$ such that when $\mathcal{H}$ is perturbed to $\mathcal{H}+\mathcal{E}$, then the eigenvalues $i\alpha$ and $i\beta$ coalesce at $i\omega_0$ to form an eigenvalue of $\mathcal{H}+\mathcal{E}$ of mixed sign characteristics. This multiple eigenvalue can then be removed from the imaginary axis by an arbitrarily small Hamiltonian perturbation of $\mathcal{H}+\mathcal{E}$.

We say that two purely imaginary eigenvalues $i\alpha$ and $i\beta$ of $\mathcal{H}$ are *adjacent* if $\mathcal{H}$ does not have an eigenvalue $i\gamma$ with $\min\{\alpha, \beta\} < \gamma < \max\{\alpha, \beta\}$.

Theorem 5.11 *Let $i\alpha$ and $i\beta$ be adjacent imaginary eigenvalues of a Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$ with $\alpha < \beta$. Let $f(\omega) := \sigma_{\min}(\mathcal{H} - i\omega I)$ for $\omega \in \mathbb{R}$, and let $\omega_0 \in (\alpha, \beta)$ be such that $f(\omega_0) = \max\{f(\omega) : \omega \in [\alpha, \beta]\}$. Set $\epsilon_0 := f(\omega_0)$. Suppose that the following conditions are satisfied.*

- *i) For all $\epsilon < \epsilon_0$ the connected components $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ and $\mathcal{C}_\epsilon(\mathcal{H}, i\beta)$ of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ containing the eigenvalues $i\alpha$ and $i\beta$, respectively, have either positive or negative sign characteristic.*

- *ii) If $\omega \in [\alpha, \beta]$ then $i\omega \in \mathcal{C}_{f(\omega)}(\mathcal{H}, i\alpha) \cup \mathcal{C}_{f(\omega)}(\mathcal{H}, i\beta)$.*

Then the following assertions hold.

- *a) The function f is strictly increasing in $[\alpha, \omega_0]$ and strictly decreasing in $[\omega_0, \beta]$. For $\epsilon < \epsilon_0$, we have $i\omega_0 \notin \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$, $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \cap \mathcal{C}_\epsilon(\mathcal{H}, i\beta) = \emptyset$ and $i\omega_0 \in \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) = \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta) = \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \cup \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta)$.*

- *b) Let $\lambda_{\min}(\cdot)$ be the function given in Proposition 5.4. If $i\alpha$ has positive sign characteristic and $i\beta$ has negative sign characteristic then $\lambda_{\min}(\omega) = f(\omega)$ for all $\omega \in [\alpha, \beta]$. On the other hand, if $i\alpha$ has negative sign characteristic and $i\beta$ has positive sign characteristic then $\lambda_{\min}(\omega) = -f(\omega)$ for all $\omega \in [\alpha, \beta]$. In both cases there exists a J-neutral normalized eigenvector v_0 of $J(\mathcal{H} - i\omega_0 I)$ corresponding to the eigenvalue $\lambda_{\min}(\omega_0)$.*

- *c) Suppose that the eigenvalues $i\alpha$ and $i\beta$ have opposite sign characteristic. Then for any $\delta > 0$ there is an $\eta >$ such that $\Lambda_{\epsilon_0+\delta}^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ contains the disk $\{z \in \mathbb{C} : |z - i\omega_0| \leq \eta\}$.*

- *d) Suppose that the eigenvalues $i\alpha$ and $i\beta$ have the same sign characteristic. Then for $\epsilon \geq \epsilon_0$, $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ is a connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ containing the eigenvalues $i\alpha$ and $i\beta$. If $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha)$ contains no other eigenvalues of $\mathcal{H}$ except $i\alpha$ and $i\beta$ then $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \subset i\mathbb{R}$ and has the same sign characteristic as that of $i\alpha$. Moreover, in such a case, there is a $\delta_0 > 0$ such that $\mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\alpha) \subset i\mathbb{R}$ for all $0 \leq \delta < \delta_0$.*

Proof. a) Observe that if $\epsilon < \epsilon_0 = f(\omega_0)$ then $i\omega_0 \notin \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H},\mathbb{F})$, and hence by assumption i) and Corollary 5.10 we have that $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \cap \mathcal{C}_\epsilon(\mathcal{H}, i\beta) = \emptyset$, and that $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \subset i\mathbb{R}$ and $\mathcal{C}_\epsilon(\mathcal{H}, i\beta) \subset i\mathbb{R}$. By assumption ii) it follows that $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \cup \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta)$ is a connected component of $\Lambda_{\epsilon_0}^{\mathrm{Ham}}(\mathcal{H},\mathbb{F})$ and hence $i\omega_0 \in \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) = \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta)$.

First, we show that f is strictly increasing in $[\alpha, \omega_0]$. Let $\gamma_1, \gamma_2 \in [\alpha, \omega_0]$ be such that $\gamma_1 < \gamma_2$. Then by assumption ii), we have $i\gamma_2 \in \mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\alpha) \cup \mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\beta)$. Now, suppose that $f(\gamma_2) < \epsilon_0 = f(\omega_0)$. Then, as we have just seen, $\mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\alpha) \cap \mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\beta) = \emptyset$, and hence $i\gamma_2 \in \mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\alpha) \subset i\mathbb{R}$. Let $\mathcal{E} \in \mathbb{F}^{2n,2n}$ be a Hamiltonian matrix such that $\|\mathcal{E}\|_2 = f(\gamma_2)$ and $i\gamma_2 \in \Lambda(\mathcal{H} + \mathcal{E})$. Setting $\mathcal{H}(t) := \mathcal{H} + t\mathcal{E}$, it follows that $\Lambda(\mathcal{H}(t)) \subset \Lambda_{f(\gamma_2)}^{\mathrm{Ham}}(\mathcal{H},\mathbb{F})$ for $t \in [0,1]$. Since $i\alpha \in \Lambda(\mathcal{H}(0))$ and $i\gamma_2 \in \Lambda(\mathcal{H}(1))$, by the continuity of eigenvalues it follows that $\Lambda(\mathcal{H}(t)) \cap \mathcal{C}_{f(\gamma_2)}(\mathcal{H}, i\alpha) \neq \emptyset$ for $t \in [0,1]$ and that $i\gamma_1 \in \Lambda(\mathcal{H}(t_0))$ for some $t_0 \in (0,1)$. Hence $f(\gamma_1) \leq \|t_0\mathcal{E}\|_2 < \|\mathcal{E}\|_2 = f(\gamma_2)$.

Next, suppose that $f(\gamma_2) = \epsilon_0 = f(\omega_0)$. If $\gamma_2 = \omega_0$ then there is nothing to prove. So, suppose that $\gamma_2 < \omega_0$. Then there exists $\gamma_3 \in (\gamma_2, \omega_0)$ such that $f(\gamma_3) < f(\omega_0) = \epsilon_0$. Since $\gamma_2, \gamma_3 \in [\alpha, \omega_0]$ with $\gamma_2 < \gamma_3$ and $f(\gamma_3) < \epsilon_0$, as we have just proved above, we have that $\epsilon_0 = f(\gamma_2) < f(\gamma_3)$, which is a contradiction. Hence, we conclude that f is strictly increasing in $[\alpha, \omega_0]$. By similar arguments, it follows f is strictly decreasing in $[\omega_0, \beta]$. This concludes the proof of a).

b) Note that $f(\alpha) = f(\beta) = 0$ and that for any $\omega \in [\alpha, \beta] \setminus \{\omega_0\}$ the connected components $\mathcal{C}_{f(\omega)}(\mathcal{H}, i\alpha)$ and $\mathcal{C}_{f(\omega)}(\mathcal{H}, i\beta)$ are disjoint, and

$$\begin{array}{ll} i[\alpha, \omega] \subseteq \mathcal{C}_{f(\omega)}(\mathcal{H}, i\alpha) & \text{if } \omega \in [\alpha, \omega_0), \\ i[\omega, \beta] \subseteq \mathcal{C}_{f(\omega)}(\mathcal{H}, i\beta) & \text{if } \omega \in (\omega_0, \beta]. \end{array} \tag{18}$$

Now consider the functions $\lambda_{\min}(\cdot)$ and $v(\cdot)$ given in Proposition 5.4. There exist finitely many numbers $-\infty = \gamma_0 < \gamma_1 < \ldots < \gamma_r < \gamma_{r+1} = \infty$ and signs $s_k \in \{-1, 1\}$ such that $\lambda_{\min}(\cdot)$ is analytic on (γ_k, γ_{k+1}) and $f(\omega) = s_k\, \lambda_{\min}(\omega)$ for $\omega \in (\gamma_k, \gamma_{k+1})$. Then $\mathcal{E}(\omega) = \lambda_{\min}(\omega) J v(\omega) v(\omega)^\star$ is Hamiltonian, $\|\mathcal{E}(\omega)\|_2 = f(\omega)$ and $(\mathcal{H} + \mathcal{E}(\omega))v(\omega) = i\omega v(\omega)$. Let $\omega \in (\alpha, \omega_0)$. Then by (18) the eigenvalue $i\omega$ of $\mathcal{H} + \mathcal{E}(\omega)$ lies in the connected component $\mathcal{C}_{f(\omega)}(\mathcal{H}, i\alpha)$ which has the same sign characteristic as that of $i\alpha$.

Suppose that $i\alpha$ has positive sign characteristic. Then $\mathcal{C}_{f(\omega)}(\mathcal{H}, i\alpha)$ has positive sign characteristic. Thus, $i\omega$ has positive sign characteristic and hence, $-iv(\omega)^\star J v(\omega) > 0$. Analogously we have $-iv(\omega)^\star J v(\omega) < 0$ for all $\omega \in (\omega_0, \beta]$. Now, for $\omega \in [\alpha, \beta] \setminus \{\gamma_1, \ldots, \gamma_r\}$, we have

$$-s_k\, iv^\star(\omega) J v(\omega) = s_k\, \lambda_{\min}'(\omega) = f'(\omega) \begin{cases} \geq 0 & \text{if } \omega \in [\alpha, \omega_0] \cap (\gamma_k, \gamma_{k+1}), \\ \leq 0 & \text{if } \omega \in [\omega_0, \beta] \cap (\gamma_k, \gamma_{k+1}). \end{cases}$$

The latter inequalities are consequences of a). It follows that $s_k = 1$ and hence, $f(\omega) = \lambda_{\min}(\omega)$ for all $\omega \in [\alpha, \beta]$. Our derivation of the latter identity was based on the assumption that $i\alpha$ has positive sign characteristic and $i\beta$ has negative sign characteristic. In the opposite case an analogous argument leads to the conclusion that $f(\omega) = -\lambda_{\min}(\omega)$ for all $\omega \in [\alpha, \beta]$. Since f is a continuous function, it now follows from Proposition 5.6 that there exists a J-neutral unit vector v_0 such that $J(\mathcal{H} - i\omega_0 I)v_0 = \lambda_{\min}(\omega_0)v_0$. This concludes the proof of b).

c) Let $\mu \in \mathbb{C}$ and consider $\mathcal{E} := \lambda_{\min}(\omega_0) J v_0 v_0^\star + \mathcal{G}(v_0, \mu\, v_0)$ when $\mathbb{F} = \mathbb{C}$, and $\mathcal{E} := \lambda_{\min}(\omega_0) J [v_0, \overline{v_0}][v_0, \overline{v_0}]^+ + \mathcal{G}([v_0, \overline{v_0}],\, [\mu v_0,\, \overline{\mu v_0}])$, when $\mathbb{F} = \mathbb{R}$, where $\mathcal{G}(\cdot, \cdot)$ is defined as in

Theorem 3.3. Then $\mathcal{E}$ is Hamiltonian and $(\mathcal{H}+\mathcal{E})v_0 = (i\omega_0+\mu)v_0$. Moreover, $\mathcal{E}$ is real when $\mathbb{F}=\mathbb{R}$. Furthermore, by Theorem 4.5, $\|\mathcal{G}(v_0, \mu\, v_0)\|_2 = |\mu|$ and $\|\mathcal{G}([v_0, \overline{v_0}], [\mu\, v_0, \overline{\mu v_0}])\|_2 \leq |\mu|(1+|v_0^T v_0|)$. Therefore, setting $\eta := \delta$ when $\mathbb{F}=\mathbb{C}$, and $\eta := \delta/(1+|v_0^T v_0|)$ when $\mathbb{F}=\mathbb{R}$, it follows that the disk $\{i\omega_0+\mu : \mu \in \mathbb{C} : |\mu| \leq \eta\}$ is contained in $\Lambda^{\mathrm{Ham}}_{\epsilon_0+\delta}(\mathcal{H}, \mathbb{F})$. This proves c).

d) Finally, w.l.o.g. suppose that both the eigenvalues $i\alpha$ and $i\beta$ have positive sign characteristic. Then both components $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ and $\mathcal{C}_\epsilon(\mathcal{H}, \beta)$ have positive sign characteristic for all $\epsilon < \epsilon_0$. Hence $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha) \cup \mathcal{C}_\epsilon(\mathcal{H}, i\beta) \subset i\mathbb{R}$ for all $\epsilon < \epsilon_0$. Recall that $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) = \mathcal{C}_{\epsilon_0}(\mathcal{H}, i\beta)$ is a connected component of $\Lambda^{\mathrm{Ham}}_{\epsilon_0}(\mathcal{H}, \mathbb{F})$. Since $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \cap \Lambda(\mathcal{H}) = \{i\alpha, i\beta\}$, by Theorem 5.9 the component $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha)$ has positive sign characteristic. Hence by Corollary 5.10, we have $\mathcal{C}_{\epsilon_0}(\mathcal{H}, i\alpha) \subset i\mathbb{R}$.

Note that the map $\epsilon \mapsto \Lambda^{\mathrm{Ham}}_\epsilon(\mathcal{H}, \mathbb{F})$ is continuous and monotonically increasing and that the components of $\Lambda^{\mathrm{Ham}}_\epsilon(\mathcal{H}, \mathbb{F})$ are closed connected sets. Hence there is a $\delta_0 > 0$ such that the component $\mathcal{C}_\epsilon(\mathcal{H}, i\alpha)$ remains disjoint from the rest of the components of $\Lambda^{\mathrm{Ham}}_\epsilon(\mathcal{H}, \mathbb{F})$ for all $\epsilon_0 \leq \epsilon < \epsilon_0 + \delta_0$. This shows that $\mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\alpha) \cap \Lambda(\mathcal{H}) = \{i\alpha, i\beta\}$ for all $0 \leq \delta < \delta_0$. Consequently, by Theorem 5.9, $\mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\alpha)$ has positive sign characteristic and hence $\mathcal{C}_{\epsilon_0+\delta}(\mathcal{H}, i\alpha) \subset i\mathbb{R}$ for all $0 \leq \delta < \delta_0$. This completes the proof. $\square$

Now consider the special case that all eigenvalues of a Hamiltonian matrix $\mathcal{H}$ are purely imaginary and each eigenvalue has either positive or negative sign characteristic. Then by Theorem 5.11 we conclude that a purely imaginary eigenvalue of $\mathcal{H}$ can be moved off from the imaginary axis only after the eigenvalue is made to coalesce with an imaginary eigenvalue of $\mathcal{H}$ of opposite sign characteristic. With a view to analyzing this issue further, we proceed as follows.

Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ be a Hamiltonian matrix whose eigenvalues are all purely imaginary, and define

$$\begin{aligned}\rho_{\mathbb{F}}(\mathcal{H}) \;&:=\; \inf\{\, \|\mathcal{E}\|_2 \,:\, \mathcal{E} \in \mathbb{F}^{2n,2n},\ (J\mathcal{E})^\star = J\mathcal{E}, \\ &\qquad\qquad \mathcal{H}+\mathcal{E} \text{ has a non-imaginary eigenvalue}\,\}, \\ R_{\mathbb{F}}(\mathcal{H}) \;&:=\; \inf\{\, \|\mathcal{E}\|_2 \,:\, \mathcal{E} \in \mathbb{F}^{2n,2n},\ (J\mathcal{E})^\star = J\mathcal{E}, \\ &\qquad\qquad \mathcal{H}+\mathcal{E} \text{ has a } J\text{-neutral eigenvector}\,\}\end{aligned}$$

Obviously, $\rho_{\mathbb{F}}(\mathcal{H}) \geq R_{\mathbb{F}}(\mathcal{H})$. The following result shows how to compute $\rho_{\mathbb{F}}(\mathcal{H})$ and $R_{\mathbb{F}}(\mathcal{H})$ using the singular value function $\omega \mapsto \sigma_{\min}(\mathcal{H} - i\omega),\ \omega \in \mathbb{R}$.

Theorem 5.12 *Let $\mathcal{H} \in \mathbb{F}^{2n,2n}$ be a Hamiltonian matrix whose eigenvalues are all purely imaginary, and let $f(\omega) = \sigma_{\min}(\mathcal{H} - i\omega I)$, $\omega \in \mathbb{R}$. Then the following assertions hold.*

i) If at least one eigenvalue of $\mathcal{H}$ has mixed sign characteristic then $R_{\mathbb{F}}(\mathcal{H}) = \rho_{\mathbb{F}}(\mathcal{H}) = 0$.

ii) Suppose that each eigenvalue of $\mathcal{H}$ has either positive or negative sign characteristic. Let $i\mathcal{I}_1, \ldots, i\mathcal{I}_q \subset i\mathbb{R}$ denote the closed intervals on the imaginary axis whose end points are adjacent eigenvalues of $\mathcal{H}$ with opposite sign characteristics. Then we have

$$R_{\mathbb{F}}(\mathcal{H}) = \rho_{\mathbb{F}}(\mathcal{H}) = \min_{1\leq k\leq q} \max_{\omega \in \mathcal{I}_k} f(\omega). \tag{19}$$

Consider an interval $\mathcal{I} \in \{\mathcal{I}_1, \ldots, \mathcal{I}_q\}$ satisfying

$$\min_{1\leq k\leq q} \max_{\omega \in \mathcal{I}_k} f(\omega) = \max_{\omega \in \mathcal{I}} f(\omega) = f(\omega_0), \quad \omega_0 \in \mathcal{I}. \tag{20}$$

Suppose that $i\mathcal{I}$ is given by $i\mathcal{I} = [i\alpha,\, i\beta]$. Then the claims a) and b) of Theorem 5.11 hold. For the J-neutral unit vector v_0 of claim b) in Theorem 5.11, consider the matrices

$$\begin{array}{rcl} \mathcal{E}^0 & := & \lambda_{\min}(\omega_0) J v_0 v_0^\star, \\ \mathcal{K}^0 & := & \lambda_{\min}(\omega_0) J [v_0\ \overline{v_0}][v_0, \overline{v_0}]^+, \\ \mathcal{E}_\mu & := & \mathcal{G}(v_0, \mu v_0), \\ \mathcal{K}_\mu & := & \mathcal{G}([v_0\ \overline{v_0}], [\mu v_0\ \overline{\mu v_0}]), \qquad \mu \in \mathbb{C}, \end{array} \tag{21}$$

where $\mathcal{G}(\cdot,\cdot)$ is defined as in Theorem 3.3. Then $\mathcal{E}^0$ is Hamiltonian, $\mathcal{K}^0$ is real and Hamiltonian, $(\mathcal{H} + \mathcal{E}^0)v_0 = (\mathcal{H} + \mathcal{K}^0)v_0 = i\omega_0 v_0$ and $\|\mathcal{E}^0\|_2 = \|\mathcal{K}^0\|_2 = f(\omega_0)$ For any $\mu \in \mathbb{C}$ the matrix $\mathcal{E}_\mu$ is Hamiltonian, and $(\mathcal{H} + \mathcal{E}^0 + \mathcal{E}_\mu)v_0 = (i\omega_0 + \mu)v_0$. If $\omega_0 = 0$ and $\mathcal{H}$ is real then v_0 can be chosen as a real vector. Then $\mathcal{E}^0 + \mathcal{E}_\mu$ is a real matrix for all $\mu \in \mathbb{R}$. If $\omega_0 \neq 0$ and $\mathcal{H}$ is real then for any $\mu \in \mathbb{C}$, $\mathcal{K}_\mu$ is a real Hamiltonian matrix satisfying $(\mathcal{H} + \mathcal{K}^0 + \mathcal{K}_\mu)v_0 = (i\omega_0 + \mu)v_0$.

Proof. Part i) is obvious.

For part ii), let ν denote the right hand side of (19), let $\omega_k \in \mathcal{I}_k$ be such that $f(\omega_k) = \max_{\omega\in\mathcal{I}_k} f(\omega)$ and let the numbering be such that $\omega_1 < \omega_2 < \ldots < \omega_q$. Then, for $0 \le \epsilon < \nu$ and all k we have $\epsilon < f(\omega_k)$, and hence $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) \cap \{i\omega_1, \ldots, i\omega_q\} = \emptyset$. Furthermore, by the definition of the intervals $\mathcal{I}_k$, the numbers $i\omega_k$ separate the eigenvalues of $\mathcal{H}$ of different sign characteristic. More precisely, for any k, all eigenvalues of $\mathcal{H}$ that are contained in the interval $i(\omega_{k-1}, \omega_k) \subset i\mathbb{R}$ have the same sign characteristic (here we use the notation $\omega_0 = -\infty$, $\omega_{q+1} = \infty$). Let $\mathcal{H}(t) = \mathcal{H} + t\mathcal{E}$, where $t \in \mathbb{R}$ and $\mathcal{E}$ is Hamiltonian with $\|\mathcal{E}\|_2 \le \epsilon$. Furthermore, let $t_0 = \sup\{\theta \in [0,1] \mid \Lambda(\mathcal{H}(t)) \subset i\mathbb{R} \text{ for all } t \in [0,\theta]\}$ and let $\Lambda_0 = \bigcup_{t\in[0,t_0]} \Lambda(\mathcal{H}(t))$. Suppose that $t_0 < 1$. Then by Theorem 4.3 the matrix $\mathcal{H}(t_0)$ has a J-neutral eigenvector. However, we have $\Lambda_0 \subseteq \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F})$ and hence, $\Lambda_0 \cap \{i\omega_1, \ldots, i\omega_q\} = \emptyset$. Thus, each connected component $\mathcal{C} \subset i\mathbb{R}$ of Λ_0 does not contain eigenvalues of $\mathcal{H} = \mathcal{H}(0)$ of opposite sign characteristic. Hence, each connected component $\mathcal{C}$ of Λ_0 has either positive or negative sign characteristic. This contradicts the assumption that $\mathcal{H}(t_0)$ has a J-neutral eigenvector. Thus, $t_0 = 1$. It follows that $\nu \le R_{\mathbb{F}}(\mathcal{H})$, $\nu \le \rho_{\mathbb{F}}(\mathcal{H})$ and $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) \subset i\mathbb{R}$ for all $\epsilon < \nu$. Furthermore, each connected component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{C})$, $\epsilon < \nu$, has either positive or negative sign characteristic.

Now, let ω_0 and $\mathcal{I}$ be as in (20). Since $i\mathcal{I} = [i\alpha, i\beta]$ and the eigenvalues $i\alpha$ and $i\beta$ have oppositive sign characteristic, the assumptions i) and ii) of Theorem 5.11 are automatically satisfied and hence the assertions a), b) and c) of Theorem 5.11 hold. The statements about the matrices $\mathcal{E}^0, \mathcal{E}_\mu, \mathcal{K}^0, \mathcal{K}_\mu$ imply that $R_{\mathbb{F}}(\mathcal{H}) \le \nu$ and $\rho_{\mathbb{F}}(\mathcal{H}) \le \nu$ which follows from Theorem 4.5 and Proposition 5.3. $\square$

Example 5.13 The eigenvalues $\pm 10i$ of the matrix $\mathcal{H}_1$ from Example 5.5 have mixed sign characteristics. Thus $R_{\mathbb{F}}(\mathcal{H}_1) = \rho_{\mathbb{F}}(\mathcal{H}_1) = 0$.

Example 5.14 Consider the Hamiltonian matrices

$$\mathcal{H}_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \end{bmatrix}, \qquad \mathcal{H}_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{bmatrix}.$$

Both matrices have the same spectrum $\Lambda(\mathcal{H}_k) = \{\pm i, \pm 2i\}$, $k = 3, 4$ and their eigenvalue curves $\omega \mapsto \lambda_j(J(\mathcal{H}_k - i\omega I))$ and singular value curves $\omega \mapsto \sigma_j(\mathcal{H}_k - i\omega I)$ are depicted in Figure 2.

Here the singular value curves for $\mathcal{H}_3$ and $\mathcal{H}_4$ coincide and the graphs of the functions $\omega \mapsto \sigma_{\min}(\mathcal{H}_k - i\omega I)$ and $\omega \mapsto \lambda_{\min}(\mathcal{H}_k - i\omega I)$ are depicted as thick curves. From the slopes of the λ_j-curves at their crossing points with the real axis we can again read off the sign characteristics of the eigenvalues $\pm i, \pm 2i$ and we see that for the matrix $\mathcal{H}_3$ the eigenvalues $-2i$ and $-i$ have negative sign characteristic, while the eigenvalues i and $2i$ have positive sign characteristic. Thus, the only pair of adjacent eigenvalues of $\mathcal{H}_3$ with opposite sign characteristic is $(-i, i)$. The maximum of the function $f(\omega) = \sigma_{\min}(\mathcal{H}_3 - i\omega I)$ in the corresponding interval $[-1, 1]$ is 1. Thus $R_{\mathbb{F}}(\mathcal{H}_3) = \rho_{\mathbb{F}}(\mathcal{H}_3) = 1$.

For the matrix $\mathcal{H}_4$ the eigenvalues $-2i$ and i have positive sign characteristic while the eigenvalues $-i$ and $2i$ have negative sign characteristic. The pairs of adjacent eigenvalues of $\mathcal{H}_4$ of opposite sign characteristic are $(-2i, -i)$, $(-i, i)$, $(i, 2i)$, and the maxima of the function $f(\omega) = \sigma_{\min}(\mathcal{H}_3 - i\omega I)$ in the corresponding intervals $[-2, -1]$, $[-1, 1]$, $[1, 2]$ are $\nu, 1, \nu$, respectively, where $\nu \approx 0.43$. Thus $R_{\mathbb{F}}(\mathcal{H}_4) = \rho_{\mathbb{F}}(\mathcal{H}_4) = \nu$.

In this section we have discussed the process of constructing the perturbations that move the eigenvalues off the imaginary axis. These will be used in the algorithm of the next section.

6 An algorithm to compute a bound for the distance to bounded-realness.

In this section we discuss a numerical method to approximately solve Problems **A** and **B**, i.e., to compute an upper bound for the smallest perturbation that moves all eigenvalues of a Hamiltonian matrix off the imaginary axis or outside a strip S_τ parallel to the imaginary axis. We cover both problems **A** and **B** by different choices of τ, i.e., Problem **A** is the case when $\tau = 0$.

In general it is an open problem to analytically classify the smallest perturbation that solves these two problems. Instead, we determine an upper bound for the smallest perturbation by solving small problems of size 2×2 or 4×4 in the real case. We also only discuss the special case that the Hamiltonian matrix has only purely imaginary eigenvalues. Numerically we can restrict ourselves to the latter case, because we can first use the methods in [5, 31] to compute a partial Hamiltonian Schur form of the matrix $\mathcal{H}$ as in (2), i.e., we determine an orthogonal (unitary) and symplectic matrix Q_0 such that for the transformed Hamiltonian matrix

$$Q_0^\star \mathcal{H} Q_0 = \left[\begin{array}{cc|cc} F_{11} & F_{12} & G_{11} & G_{12} \\ 0 & F_{22} & G_{21} & G_{22} \\ \hline 0 & 0 & -F_{11}^\star & 0 \\ 0 & H_{22} & -F_{12}^\star & -F_{22}^\star \end{array}\right],$$

we have that F_{11} is upper triangular in the complex case or quasi-upper triangular in the real case and contains those eigenvalues of $\mathcal{H}$ which lie (within the perturbation analysis of Hamiltonian matrices) [32] outside of the strip $S_\tau = \{z \in \mathbb{C} \mid -\tau < \Re z < \tau\}$.

By restricting the perturbations to the Hamiltonian submatrix

$$\widetilde{H}_2 = \begin{bmatrix} F_{22} & G_{22} \\ H_{22} & -F_{22}^\star \end{bmatrix}$$

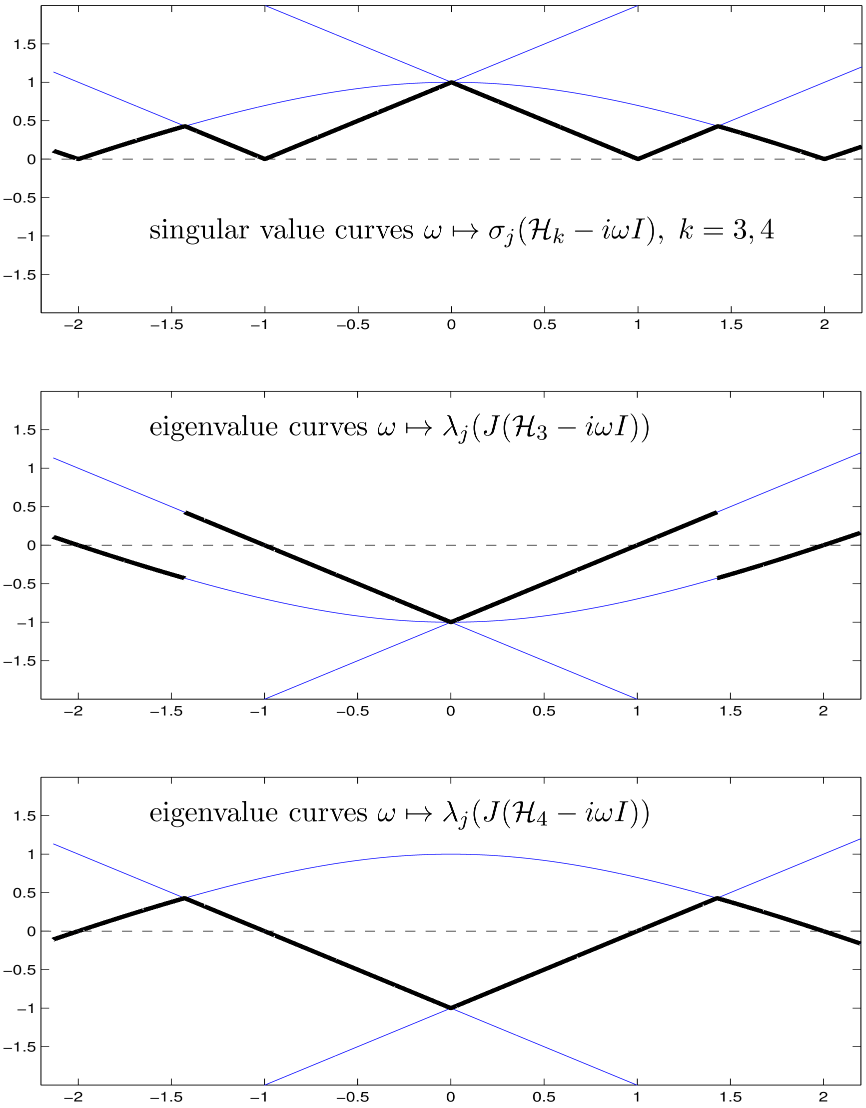

Figure 2: Eigenvalue and singular value curves for Example 5.14.

which contains all the eigenvalues that lie within the strip S_τ, we determine an upper bound for the smallest perturbation to the full matrix. The reason why it may not be the smallest perturbation is that it may be possible that the smallest perturbation first moves two eigenvalues of F_{11} that lie outside the strip S_τ into S_τ and then combines them with other eigenvalues in S_τ to get the globally smallest perturbation. We, however, do not know an example where this is the case.

There are several possibilities for the parameter τ that describes the width of the strip S_τ. It can either be preassigned to achieve a robust bounded-realness margin, or if we only want to make sure that the eigenvalues are robustly off the imaginary axis, within the usual round-off error analysis, then, since an $O(\epsilon)$ perturbation to a 2×2 Jordan block can produce an $O(\epsilon^{1/2})$ change in the eigenvalue, it seems reasonable to choose $\tau = O(u^{1/2})$, where u is the round-off unit. If there is reason to think that some of the non-imaginary eigenvalues close to the imaginary axis are the effect of round-off errors on a $k \times k$ Jordan block, then one should choose $\tau = O(u^{1/k})$.

Since, due to round-off errors, we cannot be sure whether eigenvalues of $\widetilde{H}_2$ are on or off the imaginary axis, in view of the discussed perturbation analysis we first regularize the problem by perturbing $\widetilde{H}_2$ to $\mathcal{H}_2 = \widetilde{H}_2 + \Delta H_2$ with

$$\Delta\mathcal{H}_2 = \left[\begin{array}{c|c} \Delta F_{22} & \Delta G_{22} \\ \hline \Delta H_{22} & -\Delta F_{22}^\star \end{array}\right]$$

so that all eigenvalues of $H_2 = \widetilde{H}_2 + \Delta H_2$ are on the imaginary axis. In this way the following approach, which combines nearest purely imaginary eigenvalues of opposite sign, is not restricted and we do not have to make a preliminary decision as to which eigenvalues are purely imaginary and which are not.

For each eigenvalue pair that the partial Hamiltonian Schur form produces outside the imaginary axis, a minimal perturbation $\mathcal{E}_2$ that performs this task is given by Proposition 5.3. In the following we recursively work on the matrix $\mathcal{H}_2$ and perturb one pair of purely imaginary eigenvalues at a time. Again this may have the effect of increasing the bound for the minimal perturbation, since there may be a smaller perturbation that moves several pairs at the same time.

For each chosen pair of purely imaginary eigenvalues with opposite sign characteristic (which pair of purely imaginary eigenvalues is to be chosen is discussed below) we first compute the smallest perturbation that leads to a coalescence of the pseudospectral components as described in Theorem 5.12. In this way we produce an eigenvalue of mixed sign characteristic at a point $i\gamma$. If we want to solve Problem **A**, then this perturbation is sufficient. If we want to solve Problem **B**, then we move this pair of eigenvalues to the pair $\pm\tau + i\gamma$ on the boundary of S_τ. In both cases we save the perturbation $\mathcal{E}_2$. By taking a direct sum with an appropriate 0 matrix we generate a perturbation $\mathcal{E}$ to the matrix $\mathcal{H}$ as well as its norm δ. Since in both cases the perturbed eigenvalue belongs to the part where a Hamiltonian Schur form exists, we can deflate this eigenvalue pair from $\mathcal{H}_2$ and continue with a smaller problem $\mathcal{H}_2$, for which we proceed as before.

Algorithm 1 Input: A Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$ that has only purely imaginary eigenvalues and a value $\tau > 0$ for the width of the strip S_τ around the imaginary axis.

Output: A Hamiltonian matrix $\mathcal{E} \in \mathbb{F}^{2n,2n}$, such that at least one pair (quadruple in the real case) of eigenvalues of $\mathcal{H} + \mathcal{E}$ is outside of the open strip S_τ.

Step 1: Compute the eigenvalues $i\alpha_k$, $\alpha_k \in \mathbb{R}$, $k = 1, \dots, 2n$ and associated eigenvectors $v_k \in \mathbb{C}^{2n}$ of $\mathcal{H}$. Order the eigenvalues (and eigenvectors) such that $\alpha_k \leq \alpha_{k+1}$.

Step 2: Compute the sign characteristics of the eigenvalues (i.e., the signs of $i\, v_k^\star J v_k$, $k = 1, \dots, 2n$).

Step 3: If there is a multiple eigenvalue of mixed sign characteristic, i.e., $\alpha_k = \alpha_{k+1}$, and $\mathrm{sign}(i\, v_k^\star J v_k)\, \mathrm{sign}(i\, v_{k+1}^\star J v_{k+1}) < 0$, say, then let $v_- := v_k$, $v_+ := v_{k+1}$ and go to step 6.

Step 4: For each pair of adjacent eigenvalues $i\alpha_k$, $i\alpha_{k+1}$ with opposite sign characteristic compute the maximum $m_k := \max_{\omega \in [\alpha_k, \alpha_{k+1}]} f(\omega)$, where $f(\omega) = \sigma_{\min}(\mathcal{H} - i\omega\, I) = |\lambda_{\min}(J(\mathcal{H} - i\omega I))|$, $\omega \in \mathbb{R}$.

Remark: Since f satisfies $|f(\omega) - f(\tilde{\omega})| \leq |\omega - \tilde{\omega}|$ the maxima can be found by evaluating f on a coarse grid.

Step 5: From the eigenvalues found in Step 4 select an eigenvalue $i\alpha_{k_0}$ such that $m_{k_0} = \min m_k$. By Theorems 5.11 and 5.12 there is an $\omega_0 \in [\alpha_{k_0}, \alpha_{k_0+1}]$ such that the function f is strictly increasing in $[\alpha_{k_0}, \omega_0]$ and strictly decreasing in $[\omega_0, \alpha_{k_0+1}]$ (hence $f(\omega_0) = m_{k_0}$). By using a trisection method, determine a small interval $[\omega_-, \omega_+]$ that contains ω_0. Let $v_\pm$ be eigenvectors to the eigenvalues $\lambda_{\min}(J(\mathcal{H} - i\omega_\pm\, I))$. The real numbers $-iv_\pm^\star J v_\pm$ are the slopes of the curve $\omega \mapsto \lambda_{\min}(\omega) := \lambda_{\min}(J(\mathcal{H} - i\omega\, I))$ at $\omega = \omega_\pm$. Again by Theorems 5.11 and 5.12 either $f(\omega) = \lambda_{\min}(\omega)$ for all $\omega \in [\alpha_{k_0}, \alpha_{k_0+1}]$ or $f(\omega) = -\lambda_{\min}(\omega)$ for all $\omega \in [\alpha_{k_0}, \alpha_{k_0+1}]$. Thus, $\mathrm{sign}(iv_+^\star J v_+)\, \mathrm{sign}(iv_-^\star J v_-) < 0$.

Step 6: Compute $t \in [0, 1]$ such that $u_t^\star J u_t = 0$, where $u_t = tv_+ + (1-t)v_-$, and let $v_0 = u_t / \|u_t\|$. Then v_0 is an approximate J-neutral eigenvector to the eigenvalue $\lambda_{\min}(J(\mathcal{H} - i\omega_0\, I))$.

Step 7: Let $\mu = \tau$.

Step 8: Let $\tilde{\mathcal{E}} = \mathcal{E}^0 + \mathcal{E}_\mu$ in the complex case, and $\tilde{\mathcal{E}} = \mathcal{K}^0 + \mathcal{K}_\mu$ in the real case, where $\mathcal{E}^0, \mathcal{E}_\mu, \mathcal{K}^0, \mathcal{K}_\mu$ are defined by (21). Then by Theorem 5.12, $\mathcal{H} + \tilde{\mathcal{E}}$ has (approximately) the two eigenvalues $i\omega_0 \pm \mu$ in the complex case, and the four eigenvalues $\pm i\omega_0 \pm \mu$ in the real case. Due to rounding errors $\tilde{\mathcal{E}}$ may have a slight departure from being Hamiltonian. A Hamiltonian matrix close to $\tilde{\mathcal{E}}$ is $\mathcal{E} = -\frac{1}{2} J(J\tilde{\mathcal{E}} + (J\tilde{\mathcal{E}})^\top)$.

Step 9: Check whether at least two eigenvalues of $\mathcal{H} + \mathcal{E}$ are outside the strip S_τ. If this is not the case increase μ and return to step 8.

Applying this algorithm recursively we obtain (as a sum of all the single perturbation matrices) a perturbation matrix $\Delta\mathcal{H}$ such that, at least in theory, all eigenvalues of the perturbed Hamiltonian matrix $\mathcal{H} \leftarrow \mathcal{H} + \Delta\mathcal{H}$ lie outside the strip S_τ. Due to round-off errors in the computations, however, it may happen that some eigenvalues of $\mathcal{H}$ have moved back towards the imaginary axis. Therefore, as in Step 9, it is advisable to check the spectrum of $\mathcal{H}$ to see whether the eigenvalues are safely removed from the imaginary axis in the sense that a Hamiltonian perturbation up to the size of round-off error cannot move the eigenvalues back to the imaginary axis.

So, suppose that $\mathcal{H}$ is the Hamiltonian matrix obtained by a successive application of Algorithm 1 until all eigenvalues have been moved off the imaginary axis. Then for a given tolerance τ we would like to test that the eigenvalues of $\mathcal{H}$ are robustly away from the imaginary axis in the sense that $\mathcal{H} + \mathcal{E}$ does not have an imaginary eigenvalue for any Hamiltonian perturbation $\mathcal{E}$ such that $\|\mathcal{E}\|_2 \leq \tau$. Given a Hamiltonian matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$, define

$$\beta_{\mathbb{F}}(\mathcal{H}) := \min\{\|\mathcal{E}\|_2 : \mathcal{E} \in \mathbb{F}^{2n,2n}, (J\mathcal{E})^\star = J\mathcal{E} \text{ and } \Lambda(\mathcal{H} + \mathcal{E}) \cap i\mathbb{R} \neq \emptyset\}.$$

Then $\beta_{\mathbb{F}}(\mathcal{H})$ is the distance from $\mathcal{H}$ to the Hamiltonian matrices having a purely imaginary eigenvalue. Moreover, it follows from Corollary 5.7 that $\beta_{\mathbb{F}}(\mathcal{H}) = \min\{\epsilon : \Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H}, \mathbb{F}) \cap i\mathbb{R} \neq$

$\emptyset\} = \Lambda_\epsilon(\mathcal{H},\mathbb{C}) \cap i\mathbb{R} \neq \emptyset\}$. This shows that $\beta_{\mathbb{F}}(\mathcal{H})$ is the same for $\mathbb{F}=\mathbb{R}$ and $\mathbb{F}=\mathbb{C}$ and that it can be read off from the unstructured pseudospectrum $\Lambda_\epsilon(\mathcal{H},\mathbb{C})$ of $\mathcal{H}$.

For the Hamiltonian matrix $\mathcal{H}$ computed by this procedure, we need to test whether or not $\beta_{\mathbb{F}}(\mathcal{H}) > \tau$. This can be done by computing the Hamiltonian pseudospectrum $\Lambda_\tau(\mathcal{H},\mathbb{C})$ with the method of [19] and testing whether or not $\Lambda_\tau(\mathcal{H},\mathbb{C}) \cap i\mathbb{R} = \emptyset$. Alternatively, we compute the eigenvalues of $\mathcal{H} - \tau J$ and $\mathcal{H} + \tau J$. If these matrices do not have a purely imaginary eigenvalue then by [Theorem 2, [4]] we have $\beta_{\mathbb{F}}(\mathcal{H}) > \tau$ and hence the eigenvalues of $\mathcal{H}$ are robustly away from the imaginary axis.

The computational costs of Algorithm 1 can be significantly reduced by modifying the choice of the nearest purely imaginary eigenvalues that are brought to coalescence using the following idea which may, however, in some rare cases, lead to a larger perturbation than necessary. To choose the pair $(i\gamma_1, i\gamma_2)$ or in the real case a quadruple $(i\gamma_1, -i\gamma_1, i\gamma_2, -i\gamma_2)$ of purely imaginary eigenvalues that are moved together at a point $\pm\tau + i\gamma$ we may proceed as follows. Assuming that the eigenvalues of $\mathcal{H}$ are all simple, we choose a pair of purely imaginary eigenvalue $(i\gamma_j, i\gamma_l)$ of opposite sign characteristic for which the ratio

$$\frac{|\gamma_j - \gamma_l|}{\kappa(\gamma_j) + \kappa(\gamma_l)} \tag{22}$$

is the smallest among all such pairs, where $\kappa(\gamma_j)$ is the condition number of the eigenvalue $i\gamma_j$. We arrive at this choice from the first order perturbation analysis of the eigenvalues. Indeed, by first order perturbation of eigenvalues, it follows that the component of $\Lambda_\epsilon^{\mathrm{Ham}}(\mathcal{H},\mathbb{F})$ containing $i\gamma_j$ and $i\gamma_l$ are approximately the intervals $i[\gamma_j - \kappa(\lambda_j)\epsilon,\ \gamma_j + \kappa(\gamma_j)\epsilon]$ and $i[\gamma_l - \kappa(\lambda_l)\epsilon,\ \gamma_l + \kappa(\gamma_l)\epsilon]$, respectively, for all small ϵ. Therefore, if the ratio (22) is the smallest, as ϵ increases gradually these two components are likely to coalesce before the other components.

6.1 A numerical example

To illustrate our procedure, we apply Algorithm 1 to the matrix

$$\mathcal{H} = \begin{bmatrix} -73 & -86 & 54 & -99 & 93 & -58 & 80 & 77 \\ 1 & -4 & 59 & 54 & -58 & -61 & 4 & 1 \\ -24 & -31 & -4 & -86 & 80 & 4 & 27 & 26 \\ -26 & -24 & 1 & -73 & 77 & 1 & 26 & 24 \\ -24 & -26 & -1 & -77 & 73 & -1 & 24 & 26 \\ -26 & -27 & -4 & -80 & 86 & 4 & 31 & 24 \\ -1 & -4 & 61 & 58 & -54 & -59 & 4 & -1 \\ -77 & -80 & 58 & -93 & 99 & -54 & 86 & 73 \end{bmatrix}.$$

The matrix $\mathcal{H}$ has the purely imaginary spectrum

$$\Lambda(\mathcal{H}) = \{\pm 4i, \pm 10i, \pm 16i, \pm 18i\}.$$

The intervals bounded by adjacent eigenvalues with opposite sign characteristic are $i\mathcal{I}_1 = [-16i, -10i]$, $i\mathcal{I}_2 = [-10i, -4i]$, $i\mathcal{I}_3 = [-4i, 4i]$, $i\mathcal{I}_4 = [4i, 10i]$, $i\mathcal{I}_5 = [10i, 16i]$.

Algorithm 1 computes the maximum of the function $\omega \mapsto f(\omega) = |\lambda_{\min}(\mathcal{H} - i\omega I)|$ in each of the intervals $\mathcal{I}_k$. The minimum of these maxima is attained in the interval $\mathcal{I}_1$ at

$\omega_0 \approx -13.9356$. A corresponding normalized J-neutral eigenvector (see Step 6) is

$$v_0 = \begin{bmatrix} 0.5854 - 0.2940i \\ -0.1559 - 0.1188i \\ -0.1238 - 0.0445i \\ -0.1145 - 0.0459i \\ -0.1081 - 0.0593i \\ -0.1130 - 0.0673i \\ -0.1907 - 0.0449i \\ -0.5988 - 0.2655i \end{bmatrix}.$$

For the width of the strip S_τ we choose $\tau = 0.1$. Then the output of the algorithm is the matrix (for layout reasons displayed only with 3 digits)

$$\mathcal{E} = 10^{-2} * \begin{bmatrix} 5.74 & 3.38 & 0.81 & 0.02 & 2.46 & 2.68 & 0.81 & 4.30 \\ 3.78 & 5.26 & -0.21 & -0.93 & 2.68 & 3.49 & -4.74 & 0.10 \\ 0.61 & -3.70 & -2.21 & -1.40 & -0.81 & -4.74 & 7.39 & 5.17 \\ 3.88 & -1.13 & -1.01 & -3.48 & 4.30 & 0.10 & 5.17 & 7.27 \\ -2.93 & -0.61 & -0.72 & 3.19 & -5.74 & -3.78 & -0.61 & -3.88 \\ -0.61 & -1.55 & -2.75 & 0.19 & -3.38 & -5.26 & 3.70 & 1.33 \\ -0.72 & -2.75 & 2.35 & 1.88 & -0.81 & 0.21 & 2.21 & 1.01 \\ 3.19 & 0.19 & 1.88 & 1.88 & -0.02 & 0.93 & 1.40 & 3.48 \end{bmatrix}.$$

The eigenvalues of $\mathcal{H} + \mathcal{E}$ are

$$\Lambda(\mathcal{H} + \mathcal{E}) \approx \{0.1000 \pm 13.9356i, -(0.1000 \pm 13.9356i), \pm 17.6162i, \pm 4.3627i\}.$$

A Hamiltonian Schur decomposition of $\mathcal{H} + \mathcal{E}$ yields

$$Q_0^\star(\mathcal{H} + \mathcal{E})Q_0 = \left[\begin{array}{cc|cc} F_{11} & F_{12} & G_{11} & G_{12} \\ 0 & F_{22} & G_{21} & G_{22} \\ \hline 0 & 0 & -F_{11}^\star & 0 \\ 0 & H_{22} & -F_{12}^\star & -F_{22}^\star \end{array}\right],$$

where Q_0 is symplectic and orthogonal, and

$$F_{22} = \begin{bmatrix} 7.7958 & -5.9178 \\ 7.3945 & -3.3404 \end{bmatrix}, \quad G_{22} = \begin{bmatrix} -30.8492 & -2.5331 \\ -2.5331 & 0.8874 \end{bmatrix},$$
$$H_{22} = \begin{bmatrix} 11.0658 & -5.5371 \\ -5.5371 & -0.5170 \end{bmatrix}.$$

These blocks correspond to the purely imaginary eigenvalues of $\mathcal{H}+\mathcal{E}$. By applying Algorithm 1 again to the matrix $\tilde{\mathcal{H}} = \left[\begin{array}{c|c} F_{22} & G_{22} \\ \hline H_{22} & -F_{22}^\star \end{array}\right]$ we obtain the output

$$\tilde{\mathcal{E}} = \begin{bmatrix} 0.0707 & 1.2227 & 0.7015 & 0.0862 \\ 1.2227 & 0.0306 & 0.0862 & 0.6986 \\ -2.1346 & 0.0862 & -0.0707 & -1.2227 \\ 0.0862 & -2.1375 & -1.2227 & -0.0306 \end{bmatrix}.$$

The computed eigenvalues of $\tilde{\mathcal{H}} + \tilde{\mathcal{E}}$ are

$$\Lambda(\tilde{\mathcal{H}} + \tilde{\mathcal{E}}) = \{0.1000 \pm 10.7368i, -(0.1000 \pm 10.7368i)\}.$$

Thus all eigenvalues of $\tilde{\mathcal{H}} + \tilde{\mathcal{E}}$ are outside of the open strip S_τ. Hence, there is a real Hamiltonian matrix $\Delta\mathcal{H}$ with norm

$$\|\Delta\mathcal{H}\|_2 \leq \|\mathcal{E}\|_2 + \|\tilde{\mathcal{E}}\|_2 \approx 3.005$$

such that all eigenvalues of $\mathcal{H} + \Delta\mathcal{H}$ are outside of S_τ.

7 Conclusion

We have presented a detailed perturbation analysis for eigenvalues of Hamiltonian matrices and discussed the construction of structured perturbations to Hamiltonian matrices that move eigenvalues off the imaginary axis and thereby discussed the computation of upper bounds for the distance to (robust) bounded-realness. The application of this new approach in the context of passivation problems will be discussed in forthcoming work.

References

[1] B. Adhikari and R. Alam. Structured mapping problems for a class of linearly structured matrices. Preprint, IIT Guwahati, Dept. of Mathematics, 2009.

[2] T. Antoulas. *Approximation of Large-Scale Dynamical Systems.* SIAM Publications, Philadelphia, PA, USA, 2005.

[3] A. Bunse-Gerstner. Matrix factorization for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77, 1986.

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

[5] D. Chu, X. Liu, and V. Mehrmann. A numerical method for computing the hamiltonian schur form. *Numer. Math.*, 105:375–412, 2007.

[6] C.P. Coelho, J.R. Phillips, and L.M. Silveira. Robust rational function approximation algorithm for model generation. In *Proceedings of the 36th DAC*, pages 207–212, New Orleans, Louisiana, USA, 1999.

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

[8] R.W. Freund and F. Jarre. An extension of the positive real lemma to descriptor systems. *Optimization methods and software*, 19:69–87, 2004.

[9] R.W. Freund, F. Jarre, and C. Vogelbusch. Nonlinear semidefinite programming: Sensitivity, convergence and an application in passive reduced-order modeling. *Math. Programming*, 109:581–611, 2007.

[10] P. Fuhrmann. On Hamiltonian rational transfer functions. *Linear Algebra Appl.*, 63:1–93, 1984.

[11] I. Gohberg, P. Lancaster, and L. Rodman. Spectral analysis of self-adjoint matrix polynomials. *Ann. of Math.*, 112:33–71, 1980.

[12] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications.* Birkhäuser, Basel, 2005.

[13] G.H. Golub and C.F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, third edition, 1996.

[14] S. Grivet-Talocia. Enforcing passivity of macromodels via spectral perturbation of hamiltonian matrices. In *7th IEEE Workshop on Signal Propagation on Interconnects*, pages 33–36, Siena, Italy, 2003.

[15] S. Grivet-Talocia. Passivity enforcement via perturbation of hamiltonian matrices. *IEEE Trans. Circuits Systems*, 51:1755–1769, 2004.

[16] B. Gustavsen and A. Semlyen. Enforcing passivity for admittance matrices approximated by rational functions. *IEEE Trans. on Power Systems*, 16:97–104, 2001.

[17] N. Ida and P. A. Bastos. *Electromagnetics and calculation of fields.* Springer, Verlag, New York, 1997.

[18] M. Karow. μ values and spectral value sets for linear perturbation classes defined by a scalar product. Preprint 406, Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG, 2007.

[19] M. Karow, E. Kokiopoulou, and D. Kressner. On the computation of structured singular values and pseudospectra. *Systems Control Lett.*, 59:122–129, 2010.

[20] T. Kato. *Perturbation Theory for Linear Operators.* Springer-Verlag, New York, NY, 1966.

[21] M.M. Konstantinov, V. Mehrmann, and P.Hr. Petkov. Perturbation analysis for the Hamiltonian Schur form. *SIAM J. Matrix Anal. Appl.*, 23:387–424, 2002.

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

[23] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra Appl.*, 406:1–76, 2005.

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

[25] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 301–303:469–533, 1999.

[26] C. Mehl. *Compatible Lie and Jordan algebras and applications to structured matrices and pencils.* Logos-Verlag, Berlin, 1999.

[27] C. Mehl. Condensed forms for skew-hamiltonian/hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 21:454–476, 1999.

[28] C. Mehl, V. Mehrmann, A. Ran, and L. Rodman. Eigenvalue perturbation theory of structured matrices under generic structured rank one perturbations: General results and complex matrices. Technical Report 673, Matheon, DFG Research Center *Mathematics for key technologies* in Berlin, TU Berlin, Berlin, Germany, 2009.

[29] C. Mehl, V. Mehrmann, A. Ran, and L. Rodman. Perturbation theory of selfadjoint matrices and sign characteristics under generic structured rank one perturbations. Technical Report 675, Matheon, DFG Research Center *Mathematics for key technologies* in Berlin, TU Berlin, Berlin, Germany, 2009. To appear in lin. Alg. Appl. 2010.

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

[31] V. Mehrmann, C. Schröder, and D.S. Watkins. A new block method for computing the Hamiltonian Schur form. *Linear Algebra Appl.*, 431:350–368, 2009.

[32] V. Mehrmann and H. Xu. Perturbation of purely imaginary eigenvalues of hamiltonian matrices under structured perturbations. *Electr. J. Linear Alg.*, 17:234–257, 2008.

[33] M.L. Overton and P. Van Dooren. On computing the complex passivity radius. In *Proc. 4th IEEE Conference on Decision and Control (ECC-CDC'05)*, pages 7960–7964, Seville, Spain, Dec 2005.

[34] A.C.M. Ran and L. Rodman. Stability of invariant maximal semidefinite subspaces. *Linear Algebra Appl.*, 62:51–86, 1984.

[35] A.C.M. Ran and L. Rodman. Stability of invariant Lagrangian subspaces i. *Operator Theory: Advances and Applications (I. Gohberg ed.)*, 32:181–218, 1988.

[36] A.C.M. Ran and L. Rodman. Stability of invariant Lagrangian subspaces ii. *Operator Theory: Advances and Applications (H.Dym, S. Goldberg, M.A. Kaashoek and P. Lancaster eds.)*, 40:391–425, 1989.

[37] F. Rellich. *Perturbation theory of eigenvalue problems.* Gordon and Breach Science Publishers, New York-London-Paris, 1969.

[38] D. Saraswat, R. Achar, and M. Nakhia. Enforcing passivity for rational function based macromodels of tabulated data. *Electrical Performance of Electronic Packaging*, 16:295–298, 2003.

[39] D. Saraswat, R. Achar, and M. Nakhia. On passivity check and compensation of macromodels from tabulated data. In *7th IEEE Workshop on Signal Propagation on Interconnects*, pages 25–28, Siena, Italy, 2003.

[40] C. Schröder and T. Stykel. Passivation of LTI systems. Technical Report 368, DFG Research Center Matheon, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 2007.

[41] G.W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.

[42] R. C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.

[43] L.N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators.* Princeton Univ. Press, Princeton, N.J., 2005.