Perturbation of Purely Imaginary Eigenvalues of Hamiltonian Matrices under Structured Perturbations

Volker Mehrmann+ Hongguo Xu†

October 1, 2007

Abstract

We discuss the perturbation theory for purely imaginary eigenvalues of Hamiltonian matrices under Hamiltonian and non-Hamiltonian perturbations. We show that there is a substantial difference in the behavior under these perturbations. We also discuss the perturbation of real eigenvalues of real skew-Hamiltonian matrices under structured perturbations and use these results to analyze the properties of the URV method of computing the eigenvalues of Hamiltonian matrices.

Keywords Hamiltonian matrix, skew-Hamiltonian matrix, symplectic matrix, structured perturbation, invariant subspace, purely imaginary eigenvalues, passive system, robust control, gyroscopic system

AMS subject classification. 15A18, 15A57, 65F15, 65F35.

1 Introduction

In this paper we discuss the perturbation theory for eigenvalues of Hamiltonian matrices. Let $\mathbb{F}$ denote the real or complex field and let $^*$ denote the conjugate transpose if $\mathbb{F} = \mathbb{C}$ and the transpose if $\mathbb{F} = \mathbb{R}$. Let furthermore, $J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. A matrix $\mathcal{H} \in \mathbb{F}^{2n,2n}$ is called Hamiltonian if $(J_n \mathcal{H})^* = J_n \mathcal{H}$.

The spectrum of a Hamiltonian matrix has so-called Hamiltonian symmetry, i.e., the eigenvalues appear in $(\lambda, -\bar{\lambda})$ pairs if $\mathbb{F} = \mathbb{C}$ and in quadruples $(\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda})$ if $\mathbb{F} = \mathbb{R}$.

+Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG. mehrmann@math.tu-berlin.de. Partially supported by Deutsche Forschungsgemeinschaft, through the DFG Research Center MATHEON Mathematics for Key Technologies in Berlin.

†Department of Mathematics, University of Kansas, Lawrence, KS 47045, USA. Partially supported by the University of Kansas General Research Fund allocation # 2301717 and by Deutsche Forschungsgemeinschaft through the DFG Research Center MATHEON Mathematics for Key Technologies in Berlin. Part of the work was done while this author was visiting TU Berlin whose hospitality is gratefully acknowledged.
When a given Hamiltonian matrix is perturbed to another Hamiltonian matrix, then in general the eigenvalues will change but still have the same symmetry pattern. If the perturbation is unstructured, then the perturbed eigenvalues will have lost this property.

The solution of the Hamiltonian eigenvalue problem is a key building block in many computational methods in control, see e.g. [1, 17, 30, 34, 47] and the references therein. It has also other important applications, consider the following examples.

**Example 1** *Hamiltonian matrices from robust control.* In the optimal $H_\infty$ control problem one has to deal with parameterized real Hamiltonian matrices of the form

$$H(\gamma) = \begin{bmatrix} F & G_1 - \gamma^{-2}G_2 \\ H & -F^T \end{bmatrix},$$

where $F, G_1, G_2, H \in \mathbb{R}^{n,n}$, $G_1, G_2, H$ are symmetric positive semi-definite, and $\gamma > 0$ is a parameter, see e.g., [15, 26, 45, 47]. In the $\gamma$ iteration one has to determine the smallest possible $\gamma$ such that the Hamiltonian matrix $H(\gamma)$ has no purely imaginary eigenvalues and it is essential that this $\gamma$ is computed accurately, because the optimal controller is implemented with this $\gamma$.

**Example 2** *Linear second order gyroscopic systems.* The stability of linear second order gyroscopic systems, see [20, 25, 44], can be analyzed via the following quadratic eigenvalue problem

$$P(\lambda)x = (\lambda^2I + \lambda(2\delta G) - K)x = 0,$$  \hspace{1cm} (1)

where $G, K \in \mathbb{C}^{n,n}$, $K$ is Hermitian positive definite, $G$ is nonsingular skew-Hermitian, and $\delta > 0$ is a parameter. To stabilize the system one needs to find the smallest real $\delta$ such that all the eigenvalues of $P(\lambda)$ are purely imaginary, which means that the gyroscopic system is stable.

The quadratic eigenvalue problem (1) can be reformulated as the linear Hamiltonian eigenvalue problem

$$(\lambda I - \mathcal{H}(\delta)) \begin{bmatrix} (\lambda I + \delta G)x \\ x \end{bmatrix} = 0,$$  \hspace{1cm} with $\mathcal{H}(\delta) = \begin{bmatrix} -\delta G & K + \delta^2 G^2 \\ I_n & -\delta G \end{bmatrix}$,

i.e. the stabilization problem is equivalent to determining the smallest $\delta$ such that all the eigenvalues of $\mathcal{H}(\delta)$ are purely imaginary.

A third applications arises in the context of making non-passive dynamical systems passive.

**Example 3** *Passivity of linear systems.* Consider a control system

$$\dot{x} = Ax + Bu, \ x(0) = x_0,$$

$$y = Cx + Du,$$  \hspace{1cm} (2)
with real or complex matrices $A \in \mathbb{F}^{n,n}$, $B \in \mathbb{F}^{n,m}$, $C \in \mathbb{F}^{p,n}$, $D \in \mathbb{F}^{p,m}$ and suppose that the homogeneous system is asymptotically stable, i.e. all eigenvalues of $A$ are in the open left half complex plane. The system is called \textit{passive} if it does not generate energy, i.e. for the transfer function

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

all singular values of $G(i\omega)$ are bounded from above by 1 for all $\omega \in \mathbb{R}$, [1]. It is an important task in applications from power systems [7, 18, 39] to modify a system that is non-passive to become passive. If the matrix $D + D^T$ is invertible, then the system is passive iff the Hamiltonian matrix

$$\mathcal{H} = \begin{bmatrix}
A - B(D + D^T)^{-1}C & -B(D + D^T)^{-1}B^T \\
-C^T(D + D^T)^{-1}C & -(A - B(D + D^T)^{-1}C)^T
\end{bmatrix}$$  \hspace{1cm} (3)$$

has no purely imaginary eigenvalues. Thus to make the system passive one constructs small perturbations to $A, B, C, D$ such that the system becomes passive [7, 11, 16, 39, 40].

In all these applications the location of the eigenvalues (in particular of the purely imaginary eigenvalues) of Hamiltonian matrices needs to be checked numerically at different values of parameters or perturbations. Using backward error analysis [19, 46], in finite precision arithmetic the computed eigenvalues may be considered as the exact eigenvalues of a matrix slightly perturbed from the Hamiltonian matrix.

Perturbation theory then is used to analyze the relationship between the computed and the exact eigenvalues. However, classical eigenvalue perturbation theory only shows how much a perturbation in the matrix is magnified in the eigenvalue perturbations. It usually does not give information about the location of the perturbed eigenvalues. For the purely imaginary eigenvalues of a Hamiltonian matrix, an arbitrary small unstructured perturbation to the matrix can move the eigenvalues off the imaginary axis. So in general, from the computed eigenvalues it is difficult to decide whether the exact eigenvalues are purely imaginary or not. But if the perturbation matrix is Hamiltonian, and some further properties hold, then we can show that the purely imaginary eigenvalues stay on the imaginary axis, and this property makes a fundamental difference in the decision process needed in the above discussed applications.

In recent years, a lot of effort has gone into the construction of structure preserving numerical methods to compute eigenvalues, invariant subspaces and structured Schur forms for Hamiltonian matrices. Examples of such methods are the symplectic URV methods developed in [2, 3] for computing the eigenvalues of a real Hamiltonian matrix and their extension for computing the Hamiltonian Schur form [6]. These methods produce eigenvalues and Schur forms, respectively, of a perturbed Hamiltonian matrix. To
complete the evaluation of structure preserving numerical methods in finite precision arithmetic it is then necessary to carry out a structured perturbation analysis to characterize the sensitivity of the problem under structured perturbations.

This topic has recently received a lot of attention [5, 8, 21, 22, 23, 24, 29, 43]. Surprisingly, the results in several of these papers show that the structured condition numbers for eigenvalues and invariant subspaces are often the same (or only slightly different by a factor of $\sqrt{2}$) as those under unstructured perturbations.

These observations have led to the question, whether the substantial effort that is needed to construct and implement structure preserving methods is worthwhile. However, as we will show in this paper, and this is in line with previous work [10, 35, 36, 37], there is a substantial difference between the structured and unstructured perturbation results, in particular, in the perturbation of purely imaginary eigenvalues. Let us demonstrate this with an example.

**Example 4** Consider the following Hamiltonian matrix $\mathcal{H}$ and the perturbation matrix $\mathcal{E}$ given by

$$\mathcal{H} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathcal{E} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where $a, b, c, d \in \mathbb{C}$.

The matrix $\mathcal{H}$ has two purely imaginary eigenvalues $\lambda_{1,2} = \pm i$ and the perturbed matrix

$$\mathcal{H} + \mathcal{E} = \begin{bmatrix} a & 1 + b \\ -1 + c & d \end{bmatrix}.$$

has two eigenvalues

$$\tilde{\lambda}_{1,2} = \frac{1}{2} (a + d) \pm \sqrt{(a - d)^2 - 4(1 + b)(1 - c)}.$$

The difference between the exact and perturbed eigenvalues is given by

$$\tilde{\lambda}_{1,2} - \lambda_{1,2} = \frac{1}{2} (a + d) \pm \frac{(a - d)^2 + 4(c - b + bc)}{\sqrt{(a - d)^2 - 4(1 + b)(1 - c)} + 2i},$$

so, no matter how small the perturbation $\|\mathcal{E}\|$ will be, in general, the eigenvalues will move away from the imaginary axis.

However, if $\mathcal{E}$ is Hamiltonian, then $d = -\bar{a}$ and $c, b$ are real. In this case, both eigenvalues $\tilde{\lambda}_{1,2}$ are still purely imaginary when $(1+b)(1-c)-(\text{Re} a)^2 \geq 0$, which holds when $\|\mathcal{E}\|$ is small.

This example shows the different behavior of purely imaginary eigenvalues under Hamiltonian and non-Hamiltonian perturbations. We will analyze
the reason for this difference and show that it is the existence of further invariants under structure preserving similarity transformations that are associated with these eigenvalues.

The paper is organized as follows. In Section 2 we recall some eigenvalue properties for Hamiltonian matrices. In Section 3 we describe the behavior of purely imaginary eigenvalues of Hamiltonian matrices under Hamiltonian perturbations. In Section 4 we study the behavior of real eigenvalues of skew-Hamiltonian matrices with skew-Hamiltonian perturbations. In Section 5, we then derive conditions so that the symplectic URV algorithm proposed in [3] can correctly compute the purely imaginary eigenvalues of a real Hamiltonian matrix. We finish with some conclusions in Section 6.

2 Notations and Preliminaries

The subspace spanned by the columns of matrix $X$ is denoted by $\text{span } X$. $I_n$ (or simply $I$) is the identity matrix. The spectrum of a square matrix $A$ is denoted by $\lambda(A)$. The spectrum of a matrix pencil $\lambda E - A$ is denoted by $\lambda(E, A)$. $\| \cdot \|$ denotes a vector norm or a matrix norm.

We use the notation $a = O(b)$ to indicate that $|a/b| \geq C$ for some positive constant $C$ as $b \to 0$, and the notation $a = o(b)$ to indicate that $|a/b| \to 0$ as $b \to 0$.

**Definition 1** For a matrix $A \in \mathbb{F}^{n,n}$, a subspace $V \subseteq \mathbb{F}^n$ is a right (left) invariant subspace of $A$ if $AV \subseteq V$ ($A^*V \subseteq V$). Let $\lambda$ be an eigenvalue of $A$. A right invariant subspace $V$ of $A$ is the right (left) eigenspace corresponding to $\lambda$ if $A|_V$ ($A^*|_V$), the restriction of $A$ to $V$, has the single eigenvalue $\lambda$, and $\dim V$ equals the algebraic multiplicity of $\lambda$.

**Definition 2**

(i) A matrix $S \in \mathbb{F}^{2n,2n}$ is called symplectic if $S^*J_n S = J_n$.

(ii) A matrix $U \in \mathbb{C}^{2n,2n}$ ($\mathbb{R}^{2n,2n}$) is called unitary (orthogonal) symplectic if $U$ is symplectic and $U^*U = I_{2n}$.

(iii) A matrix $H \in \mathbb{F}^{2n,2n}$ is called Hamiltonian if $H J_n = (H J_n)^*$.

(iv) A matrix $K \in \mathbb{F}^{2n,2n}$ is called skew-Hamiltonian if $K J_n = -(K J_n)^*$.

By $A \otimes B = [a_{ij} B]$ we denote the Kronecker product of $A$ and $B$. We introduce the particular matrices

$$P_r = \begin{bmatrix} 1 & & & \cdots & 1 \\ 1 & & & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & & & \cdots & 1 \end{bmatrix}, \quad \hat{P}_r = \begin{bmatrix} 1 & & & \cdots & (1)^0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & & & \cdots & (-1)^{r-1} \end{bmatrix}, \quad N_r = \begin{bmatrix} 0 & 1 & & & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots \\ & & & & 1 \end{bmatrix},$$
and
\[ N_r(a) = aI_r + N_r, \quad N_r(a, b) = I_r \otimes \begin{bmatrix} a & b \\ -b & a \end{bmatrix} + N_r \otimes I_2. \]

For Hamiltonian and skew-Hamiltonian matrices structured Jordan canonical forms are well known.

**Theorem 2.1 ([12, 27, 28, 42])** For a Hamiltonian matrix \( H \in \mathbb{C}^{2n, 2n} \) there exists a nonsingular matrix \( X \) such that
\[ X^{-1}HX = \text{diag}(H_1, \ldots, H_m), \quad X^*J_nX = \text{diag}(Z_1, \ldots, Z_m), \]
where each pair \((H_j, Z_j)\) is of one of the following forms:

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

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

The scalars \( s_j \) in Theorem 2.1 are called the sign characteristic of the pair \((H, J)\) associated with the purely imaginary eigenvalues, they satisfy \( \sum s_j = 0 \).

In the real case the canonical form is as follows.

**Theorem 2.2 ([12, 27, 28, 42])** For a Hamiltonian matrix \( H \in \mathbb{R}^{2n, 2n} \) there exists a real nonsingular matrix \( X \) such that
\[ X^{-1}HX = \text{diag}(H_1, \ldots, H_m), \quad X^TJ_nX = \text{diag}(Z_1, \ldots, Z_m), \]
where each pair \((H_j, Z_j)\) is of one of the following forms:

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

(a.2) \( H_j = \begin{bmatrix} N_{2n_j+1} & 0 \\ 0 & -N_{2n_j+1}^T \end{bmatrix}, \quad Z_j = \begin{bmatrix} 0 & I_{2n_j+1} \\ -I_{2n_j+1} & 0 \end{bmatrix} = J_{2n_j+1}, \)

corresponding to two \((2n_j+1) \times (2n_j+1)\) Jordan blocks for the eigenvalue 0,

(b) \( H_j = \begin{bmatrix} 0 & N_{n_j}(\alpha_j) \\ -N_{n_j}(\alpha_j) & 0 \end{bmatrix}, \quad Z_j = s_j \begin{bmatrix} 0 & P_{n_j} \\ -P_{n_j} & 0 \end{bmatrix}, \) where \( 0 < \alpha_j \in \mathbb{R} \) and \( s_j = \pm 1 \), corresponding to an \( n_j \times n_j \) Jordan block for each of the purely imaginary eigenvalues \( \pm \alpha_j \).
\( H_j = \begin{bmatrix} N_{n_j}(\beta_j) & 0 \\ 0 & -[N_{n_j}(\beta_j)]^T \end{bmatrix}, \quad Z_j = \begin{bmatrix} 0 & I_{n_j} \\ -I_{n_j} & 0 \end{bmatrix} = J_{n_j} \) where \( 0 < \beta_j \in \mathbb{R}, \) corresponding to an \( n_j \times n_j \) Jordan block for each of the real eigenvalues \( \beta_j \) and \(-\beta_j\),

\( H_j = \begin{bmatrix} N_{n_j}(a_j, b_j) & 0 \\ 0 & -[N_{n_j}(a_j, b_j)]^T \end{bmatrix}, \quad Z_j = \begin{bmatrix} 0 & I_{2n_j} \\ -I_{2n_j} & 0 \end{bmatrix} = J_{2n_j} \) 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 \( t_j, s_j \) of the pair \((\mathcal{H}, J)\). Note further that both in the real and complex Hamiltonian cases the transformation matrix \( X \) can be constructed to be a symplectic matrix, see [28].

For complex skew-Hamiltonian matrices the canonical form is similar to the Hamiltonian canonical form, since if \( \mathcal{K} \) is skew-Hamiltonian then \( i\mathcal{K} \) is Hamiltonian. For real skew-Hamiltonian matrices the canonical form is different but simpler.

**Theorem 2.3 ([9, 42])** *For a skew Hamiltonian matrix \( \mathcal{K} \in \mathbb{R}^{2n,2n} \) there exists a real symplectic matrix \( S \) such that*

\[
S^{-1}\mathcal{K}S = \begin{bmatrix} K & 0 \\ 0 & K^T \end{bmatrix},
\]

*where \( K \) is in real Jordan canonical form.*

Theorem 2.3 shows that every Jordan block of \( \mathcal{K} \) appears twice, and thus the algebraic and geometric multiplicity of every eigenvalue must be even.

After introducing some notation and recalling the canonical forms, in the next section we study the perturbation of purely imaginary eigenvalues of Hamiltonian matrices.

3 Perturbations of purely imaginary eigenvalues of Hamiltonian matrices

Let \( \mathcal{H} \in \mathbb{C}^{2n,2n} \) 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 eigenspace associated with \( i\alpha \), i.e.,

\[
\mathcal{H}X = XR, \tag{4}
\]

where \( \lambda(R) = \{i\alpha\} \). By using the Hamiltonian property \( \mathcal{H} = -J\mathcal{H}^*J^* \), we also have

\[
X^*J\mathcal{H} = -R^*X^*J. \tag{5}
\]
Since also $\lambda(-R^*) = \{i\alpha\}$, it follows that the columns of full column rank matrix $J^*X$ span the left eigenspace of $i\alpha$. Hence, we have that

$$(J^*X)^*X = X^*JX$$

is nonsingular. Then the matrix

$$Z := iX^*JX$$

is Hermitian and nonsingular. The matrix

$$M := X^*(J\mathcal{H})X,$$

is also Hermitian. Thus, by pre-multiplying $X^*J$ to (4), we obtain

$$M = Z(-iR) \Rightarrow -iR = Z^{-1}M.$$  

This implies that the spectrum of the pencil $\lambda Z - M$ is given by

$$\lambda(Z, M) = \lambda(-iR) = \{\alpha\}.$$  

We combine these observations in the following lemma.

**Lemma 1** Let $\mathcal{H} \in \mathbb{C}^{2n,2n}$ be Hamiltonian and suppose that $i\alpha$ is a purely imaginary eigenvalue of $\mathcal{H}$. Suppose that $\alpha \neq 0$ and let $R, X, Z$ be as defined in (4) and (6).

Then the matrix $Z$ is definite if and only if $i\alpha$ is a multiple eigenvalue with equal algebraic and geometric multiplicity and has uniform sign characteristic.

**Proof.** Since the Sylvester inertia index of $Z$, i.e. the number of positive, negative or zero eigenvalues, is independent of the choice of basis, based on the structured canonical form in Theorem 2.1 (a), we may choose $X$ such that

$$R = \text{diag}(iN_{n_1}(\alpha), \ldots, iN_{n_q}(\alpha)), \quad Z = -\text{diag}(s_1P_{n_1}, \ldots, s_pP_{n_q}),$$

where $s_j = \pm 1$ are the corresponding sign characteristics.

If $n_j > 1$ for some $j$, then $Z$ has a diagonal block $-s_jP_{n_j}$ which is indefinite and thus $Z$ is indefinite. If $n_1 = \ldots = n_q = 1$, $Z = -\text{diag}(s_1, \ldots, s_q)$, then $Z$ is definite if and only if all $s_j$ ($j = 1, \ldots, q$) have the same sign. 

With the help of Lemma 1 we can now characterize the behavior of purely imaginary eigenvalues under Hamiltonian perturbations. This topic is well studied in the more general context of self-adjoint matrices with respect to indefinite inner products, see [13, 14]. A detailed perturbation analysis of the sign characteristics is given in [38] for the case that the Jordan structure is kept fixed as well as the multiplicities of nearby eigenvalues. We are, in particular, interested in the purely imaginary eigenvalues of Hamiltonian matrices and to characterize how many eigenvalues stay on the imaginary axis and how many move away from the axis. This cannot be concluded directly from these results, so we present a different analysis and proof.
Theorem 3.1. Consider the Hamiltonian matrix $H \in \mathbb{C}^{2n \times 2n}$ with a purely imaginary eigenvalue $i\alpha$ of algebraic multiplicity $p$. Suppose that $X \in \mathbb{C}^{2n \times p}$ satisfies $\text{rank} \ X = p$ and (4), and that $Z$ and $M$ are defined as in (6) and (7), where $Z$ is congruent to $\begin{bmatrix} I_{\pi} & 0 \\ 0 & -I_{\mu} \end{bmatrix}$ (with $\pi + \mu = p$).

If $E$ is Hamiltonian and $\|E\|$ is sufficiently small, then $H + E$ 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 possibilities.

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
   $$\lambda_j = i(\alpha + \delta_j) + O(\|E\|),$$
   where $\delta_1, \ldots, \delta_p$ are the real eigenvalues of the pencil $\lambda Z - X^* (J E) X$.

2. If there exists a Jordan block associated with $i\alpha$ of size larger than 2, then generically for a given $E$ some eigenvalues of $H + E$ 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 $E$ with $\|E\| = \epsilon$ such that some eigenvalues of $H + E$ 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 $E$ with $\|E\| = \epsilon$ such that some eigenvalues of $H + E$ will have nonzero real part.

Proof. Consider the Hamiltonian matrix

$$H(t) = H + tE,$$

with $0 \leq t \leq 1$. If $\|E\|$ is sufficiently small, then by the classical invariant subspace perturbation theory [41], there exists a full column rank matrix $X(t) \in \mathbb{C}^{2n \times p}$, which is a continuous function of $t$ satisfying $X(0) = X$ such that

$$H(t)X(t) = X(t)R(t),$$

where $R(t)$ has the eigenvalues $\lambda_1(t), \ldots, \lambda_p(t)$. These $p$ eigenvalues satisfy $\lambda_1(0) = \ldots = \lambda_p(0) = i\alpha$, and are close to $i\alpha$ and separated from the rest of the eigenvalues of $H(t)$, for $0 < t \leq 1$.

Since $H(t)$ is Hamiltonian, similarly, for

$$Z(t) = iX(t)^* JX(t), \quad M(t) = X(t)^* (JH(t))X(t),$$

9
we have the properties that $Z(t)$ is Hermitian nonsingular, $M(t)$ is Hermitian and
\[ \lambda(Z(t), M(t)) = \lambda(-iR(t)) = \{-i\lambda_1, \ldots, -i\lambda_p\}. \]
Because $Z(t)$ is also a continuous function of $t$, and $Z(0) = Z$, $\det Z(t) \neq 0$ for $0 \leq t \leq 1$, we have that $Z(t)$ is congruent to $\begin{bmatrix} I_0 & 0 \\ 0 & -I_\mu \end{bmatrix}$, i.e., $Z(t)$ has the same inertia as $Z$ for $0 \leq t \leq 1$. Therefore $Z(1)$ is congruent to $\begin{bmatrix} I_0 & 0 \\ 0 & -I_\mu \end{bmatrix}$.

Based on the structured canonical form of Hermitian/Hermite pencils [31, 42], the pencil $\lambda Z(1) - M(1)$ has at least $|\pi - \mu|$ real eigenvalues. Since $\lambda(R(1)) = i\lambda(Z(1), M(1))$, we conclude that $R(1)$, or equivalently $\mathcal{H} + \mathcal{E} = \mathcal{H}(1)$ has at least $|\pi - \mu|$ purely imaginary eigenvalues near $i\alpha$.

When $Z$ is definite, i.e. either $\pi = 0$ or $\mu = 0$ then by the same argument as before $Z(1)$ is definite. Therefore, since $|\pi - \mu| = p$, $\mathcal{H} + \mathcal{E}$ has $p$ purely imaginary eigenvalues $\lambda_1 := \lambda_1(1), \ldots, \lambda_p := \lambda_p(1)$.

We still need to prove the first order perturbation formula in part 1., and the other two cases. It follows from (4) and (5) that
\[ X^*JXR = -R^*X^*JX, \]
i.e.,
\[ -(X^*JX)^{-1}R^* = R(X^*JX)^{-1}. \]
Then, for $Y = ((X^*JX)^{-1}X^*J)^*$, by (5) we have
\[ Y^*\mathcal{H} = RY^*, \]
and $Y^*X = I_p$.

It follows from first order perturbation theory [41], that for $\|\mathcal{E}\|$ sufficiently small, the eigenvalues $\lambda_1, \ldots, \lambda_p$ of $\mathcal{H} + \mathcal{E}$ that are close to $i\alpha$ are eigenvalues of
\[ \tilde{R} = Y^*(\mathcal{H} + \mathcal{E})X + \delta E, \]
where $|\delta E| = O(\|\mathcal{E}\|^2)$. This relation can be written as
\[ \tilde{R} = R + iZ^{-1}X^*(J\mathcal{E})X + \delta E. \] (11)
If $Z$ is definite, then the Hermitian pencil $\lambda Z - X^*(J\mathcal{E})X$ as well as the matrix $Z^{-1}X^*(J\mathcal{E})X$ have real eigenvalues $\delta_1, \ldots, \delta_p$. By Lemma 1, when $Z$ is definite, it follows that $R = i\alpha I$ and thus for the eigenvalues of the perturbed problem we have
\[ \lambda_j = i\alpha + i\delta_j + O(\|\mathcal{E}\|^2), \]
for $j = 1, \ldots, p$.

We now prove the first part of part 2. The second part will be given in Example 5 below. W.l.o.g. we assume that $R$ and $Z$ are given as in (9), where $n_j > 1$ for some $j$. Note that (11) still holds. Generically for
unstructured perturbations, the largest eigenvalue perturbation occurs in
the largest Jordan blocks, see [32] and for generic structured perturbations
of Hamiltonian matrices this is true as well [24]. Thus, for simplicity of
presentation, we assume that \( R \) consists of \( q \) equal blocks of size \( r := n_1 = \ldots = n_q > 1 \). (Otherwise, we could first only consider the perturbation of
the submatrix associated with the largest Jordan blocks.) Then with an
appropriate permutation \( P, R \) and \( Z \) can be simultaneously transformed to
\( R_1 = P^T R P \) and \( Z_1 = P^T Z P \) as

\[
R_1 = i \begin{bmatrix} \alpha I_q & I_q & 0 \\ & \ddots & \ddots \\ 0 & \cdots & \alpha I_q \end{bmatrix}_{r \times r}, \quad Z_1 = -\begin{bmatrix} 0 & S \\ S & 0 \end{bmatrix}_{r \times r},
\]

where \( S = \text{diag}(s_1, \ldots, s_q) \). Let \( \tilde{R}_1 = P^T R P, E = P^T X^* (jE) X P, \delta E_1 = P^T (\delta E) P \), where \( \tilde{R}, X^* (jE) X, \) and \( \delta E \) are given in (11). Let \( E = [E_{ij}] \) be
partitioned conformably with \( R_1 \) and \( Z_1 \). Then we have

\[
\tilde{R}_1 = i \begin{bmatrix} \alpha I - SE_{1r}^* & I - SE_{2r}^* & \ldots & -SE_{r-1,r}^* & -SE_{rr} \\ -SE_{1,r-1}^* & \alpha I - SE_{2,r-1}^* & \ldots & -SE_{r-1,r-1}^* & -SE_{r-1,r} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -SE_{12}^* & -SE_{22}^* & \ldots & \alpha I - SE_{2,r-1} & I - SE_{2r} \\ -SE_{11} & -SE_{12} & \ldots & -SE_{1,r-1} & \alpha I - SE_{1r} \end{bmatrix} + \delta E_1,
\]

where generically \( E_{11} \) is nonzero. Let \( \tilde{\lambda} \) be an eigenvalue of \( \tilde{R}_1 \) and \( \tilde{x} = [\tilde{x}_1^T, \ldots, \tilde{x}_r^T]^T \) with \( \|\tilde{x}\| = 1 \) be a corresponding eigenvector, partitioned in
the same way. By using the block structure, we have

\[
[(\tilde{\lambda} - i\alpha)^r I + i^{r-1} SE_{11}] \tilde{x}_1 = v_1,
\]

where \( v_1 \) is a vector satisfying \( \|v_1\| = o(\|E\|) \). Let \( \gamma_1, \ldots, \gamma_t \) be the nonzero
eigenvalues of \( iSE_{11} \), then \( \tilde{R}_1 \) has at least \( rt \) eigenvalues as

\[
i\alpha + i\rho_{k,j} + o(\|E\|^{1/2})
\]

for \( j = 1, \ldots, r \) and \( k = 1, \ldots, t \), where \( \rho_{k,1}, \ldots, \rho_{k,r} \) are the \( r \)-th roots of
\( \gamma_k \).

Clearly, if \( r > 2 \), then in general there always exist some non-real \( \rho_{k,j} \),
which implies that \( \tilde{R}_1 \) and, therefore, also \( \mathcal{H} + E \) have some eigenvalues
nearby \( i\alpha \) but with nonzero real parts.

Finally, to prove part 3., we may assume that \( R = i\alpha I_p \) and \( Z = \begin{bmatrix} I_r & 0 \\ 0 & -I_r \end{bmatrix} \) with \( \pi, \mu > 0 \). From (11) it follows that

\[
\tilde{R} = i\alpha I + iZ^{-1} X^* (jE) X + \delta E.
\]
In this case the eigenvalues of $\tilde{R}$ can be expressed as

$$\lambda_j = i\alpha + \rho_j + O(\|E\|^2),$$

where $\rho_1, \ldots, \rho_p$ are the eigenvalues of $iZ^{-1}X^*(J\mathcal{E})X$, or equivalently, the Hermitian/skew-Hermitian pencil

$$\lambda Z - iX^*(J\mathcal{E})X.$$ 

Since $Z$ is indefinite, one can always find $\mathcal{E}$, no matter how small $\|\mathcal{E}\|$ is, such that this pencil has eigenvalues with nonzero real part and thus also $\lambda_j$ must have nonzero real part.

**Remark 1** The eigenvalues $\delta_1, \ldots, \delta_p$ in the first order perturbation formula are independent of the choice of the subspace $X$. If $Y$ is a full column rank matrix such that span $Y = \text{span} X$, then $Y = XT$ for some nonsingular matrix $T$. Therefore,

$$\lambda iY^* JY - Y^* (J\mathcal{E})Y = T^*(\lambda Z - X^*(J\mathcal{E})X)T.$$ 

Clearly, $X$ can be chosen to be orthonormal. If $p = 1$, then for such an $X$, the associated $Z$ is the reciprocal of the condition number of $i\alpha$.

In the following we give two examples to illustrate parts 2. and 3. in Theorem 3.1.

**Example 5** Suppose that $\mathcal{H}$ consists of only one $2 \times 2$ Jordan block associated with the purely imaginary eigenvalue $i\alpha$, and let $X$ be a full column rank matrix such that

$$X^*JX = is \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathcal{H}X = X \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix} = XR,$$

where $s = \pm 1$. Let $\mathcal{E}$ be Hamiltonian and

$$E = iZ^{-1}X^*(J\mathcal{E})X = (X^*JX)^{-1}X^*(J\mathcal{E})X = -is \begin{bmatrix} \bar{b} & c \\ a & b \end{bmatrix},$$

where $a, c \in \mathbb{R}$ (because $X^*(J\mathcal{E})X = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix}$ is Hermitian). Then the eigenvalues of $\tilde{R}$ are

$$\lambda_{1,2} = i \left( \alpha - s \Re b \pm \sqrt{-(\Im b)^2 - a(s - c)} \right) + O(\|\mathcal{E}\|^2).$$

No matter how small $\|\mathcal{E}\|$ is, we can always find $\mathcal{E}$ with $(\Im b)^2 + a(s - c) > 0$. Then both $\lambda_1, \lambda_2$ have nonzero real part.
Theorem 3.1 can now be used to explain why the purely imaginary eigenvalues \( \pm i \) of the Hamiltonian matrix \( H \) in Example 4 are hard to move off the imaginary axis by Hamiltonian perturbations. The reason is that if we take the eigenvectors of \( i \) and \( -i \) as \( \begin{bmatrix} 1 \\ i \end{bmatrix} \) and \( \begin{bmatrix} 1 \\ -i \end{bmatrix} \), respectively, the corresponding matrices \( Z \) are \( -2 \) and \( 2 \), respectively, which are both definite.

Let us consider another example.

**Example 6** The Hamiltonian matrices

\[
H_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}
\]

both have a pair of purely imaginary eigenvalues \( \pm i \) with algebraic multiplicity 2. For the eigenvalue \( i \) the right eigenspaces are spanned by the columns of the matrices

\[
X_1 = \begin{bmatrix} 1 & 0 \\ i & 0 \\ 0 & 1 \\ 0 & i \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ i & 0 \\ 0 & i \end{bmatrix},
\]

and the corresponding \( Z \) matrices are

\[
Z_1 = 2i \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad Z_2 = -2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\]

respectively. For \( H_1 \) a small Hamiltonian perturbation may move \( i \) off the imaginary axis, while for \( H_2 \) only a large Hamiltonian perturbation can achieve this.

For an eigenvalue of a Hamiltonian matrix that is not purely imaginary, based on the Hamiltonian canonical form, it is not difficult to show that a general Hamiltonian perturbation (regardless of the perturbation magnitude) is very unlikely to move the eigenvalue to the imaginary axis. If one really wants to construct a perturbation that moves an eigenvalue with nonzero real part to the imaginary axis, then one needs to construct special perturbation matrices. [33].

In the context of the passivation problem in Example 3 one has to study the problem, what is the minimal norm of a Hamiltonian perturbation \( \mathcal{E} \) that is needed in order to move purely imaginary eigenvalues off the imaginary axis. If the purely imaginary eigenvalues are as in parts 2. and 3. of Theorem 3.1, then certainly a perturbation with arbitrary small norm will achieve this goal. So one only needs to consider the case that all the purely imaginary eigenvalues are as in part 1. of Theorem 3.1. We then have the following result.
Theorem 3.2 Suppose that $\mathcal{H} \in \mathbb{C}^{2n,2n}$ is Hamiltonian and all its eigenvalues are purely imaginary. Let $\mathbb{H}_{2n}$ be the set of $2n \times 2n$ complex Hamiltonian matrices, and let $\mathcal{S}$ be the set of Hamiltonian matrices defined by

$$\mathcal{S} = \left\{ \mathcal{E} \in \mathbb{H}_{2n} \mid \mathcal{H} + \mathcal{E} \text{ has an imaginary eigenvalue with algebraic multiplicity} > 1 \text{ and the corresponding } Z \text{ in (6) is indefinite} \right\}.$$

Define

$$\mu_0 = \min_{\mathcal{E} \in \mathcal{S}} \|\mathcal{E}\|.$$

If every eigenvalue of $\mathcal{H}$ has equal algebraic and geometric multiplicity and the corresponding matrix $Z$ as in (6) is definite, then for any Hamiltonian matrix $\mathcal{E}$ with $\|\mathcal{E}\| \leq \mu_0$, $\mathcal{H} + \mathcal{E}$ has only purely imaginary eigenvalues. For any $\mu > \mu_0$, there always exists a Hamiltonian matrix $\mathcal{E}$ with $\|\mathcal{E}\| = \mu$ such that $\mathcal{H} + \mathcal{E}$ has an eigenvalue with nonzero real part.

Proof. By the assumption, we may divide the purely imaginary eigenvalues into $k$ sets $L_1, \ldots, L_k$ such that

(i) for $i < j$, any eigenvalue $\lambda_{i,\ell} \in L_i$ is below all the eigenvalues $\lambda_{j,\ell} \in L_j$ on the imaginary axis,

(ii) the eigenvalues $\lambda_{i,\ell}$ in each set $L_i$ have the corresponding matrices $Z_{i,\ell}$ as in (6) either all positive definite or all negative definite, and

(iii) the definiteness is alternating on the sets, i.e., if the $Z_{j,\ell}$ associated with the set $L_j$ are all positive definite, then they are all negative definite on $L_{j+1}$.

For the eigenvalues in set $L_j$ we define the full column rank matrix $X_j$ such that the columns of $X_j$ span the corresponding invariant subspace. Also we define $Z_j = iX_j^*JX_j$ for $j = 1, \ldots, k$.

Because in the set $L_j$ for all the eigenvalues the corresponding matrices $Z_{j,\ell}$ have the same definiteness, $Z_j$ is also definite and the definiteness alternates from positive to negative.

Now let $\mathcal{E}$ be a Hamiltonian perturbation. Then the eigenvalues of $\mathcal{H} + \mathcal{E}$ are continuous functions of the entries of $\mathcal{E}$. By using the continuity argument as for case 1. of Theorem 3.1, it follows that the eigenvalues of $\mathcal{H} + \mathcal{E}$ are all purely imaginary and in $k$ separated sets $L_1(\mathcal{E}), \ldots, L_k(\mathcal{E})$, regardless of their algebraic multiplicity, until for some $\mathcal{E}$ two neighboring eigenvalue sets have a common element. The common element is an eigenvalue of $\mathcal{H} + \mathcal{E}$ and the associated matrix $Z$ must be indefinite, since its invariant subspace consists of the subspaces from those invariant subspaces corresponding to the eigenvalues from two adjacent sets. In this case, based on Theorem 3.1 any small Hamiltonian perturbation may move this eigenvalue off the imaginary axis. Using the Hamiltonian canonical form one can always find such a
perturbation $\mathcal{E}$. So the minimization problem has a well-defined minimum.

If $\mathcal{H}$ also has some eigenvalues that are not purely imaginary, then the situation is much more complicated, and in general Theorem 3.2 does not hold. The complexity of the problem in this case can be illustrated by the following example.

**Example 7** Consider a $4 \times 4$ Hamiltonian matrix $\mathcal{H}$ with two purely imaginary eigenvalues $i\alpha_1, i\alpha_2$ and eigenvalues $\lambda, -\bar{\lambda}$, with nonzero real part, see Figure 1. If we consider the perturbation $\mathcal{E}$ to move $i\alpha_1, i\alpha_2$ off the imaginary axis, while freezing $\lambda, -\bar{\lambda}$, then we basically obtain the same result as in Theorem 3.2 with an additional restriction to $\mathcal{E}$. However, the involvement of $\lambda, -\bar{\lambda}$ may help to reduce the norm of $\mathcal{E}$ to move off $i\alpha_1, i\alpha_2$. Suppose that $\lambda, -\bar{\lambda}$ are already close to the imaginary axis as in Figure 1 (1). We may construct first a small perturbation that moves $\lambda, -\bar{\lambda}$ to the imaginary axis, forming a double purely imaginary eigenvalues $i\beta_1, i\beta_2$ and by continuity it follows that its associated matrix $Z$ is indefinite. Meanwhile, we also move $i\alpha_1, i\alpha_2$ towards each other with an appropriate perturbation. Next we determine another perturbation that forces $i\beta_1, i\beta_2$ to move in opposite direction along the imaginary axis to meet $i\alpha_1$ and $i\alpha_2$, respectively, see Figure 1 (2). Also, we force the $Z$ matrices (they are scalar in this case) associated with $i\alpha_j, i\beta_j$ to have opposite signs for $j = 1, 2$. When $i\alpha_1, i\beta_1$
and $i\alpha_2, i\beta_2$ meet, they will be moved off the imaginary axis, see Figure 1 (3). It is then possible to have an $\mathcal{E}$ with a norm smaller than the one that freezes $\lambda, -\bar{\lambda}$.

To make this concrete, consider the Hamiltonian matrix

$$\mathcal{H} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & \mu & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu \end{bmatrix},$$

where $0 < \mu < \frac{1}{2}$. Then $\mathcal{H}$ has two purely imaginary eigenvalues $i, -i$ and two real eigenvalues $\mu, -\mu$. The Hamiltonian matrices of the form

$$\mathcal{E}_1 = \begin{bmatrix} a & 0 & b & 0 \\ 0 & 0 & 0 & 0 \\ c & 0 & -\bar{a} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

with $b, c \in \mathbb{R}$ only perturb the eigenvalues $i$ and $-i$. Using the eigenvalue formula obtained in Example 4 and some elementary analysis, the minimum 2-norm of $\mathcal{E}_1$ for both $i$ and $-i$ to move to $0$ is $1$ (say, with $a = 0, b = 0$ and $c = 1$). Then a random Hamiltonian perturbation with an arbitrary small 2-norm will move the eigenvalues off the imaginary axis.

On the other hand, with the Hamiltonian perturbation

$$\mathcal{E}_2 = \begin{bmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\mu & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \mu \end{bmatrix},$$

we have

$$\mathcal{H} + \mathcal{E}_2 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

This matrix has a pair of purely imaginary eigenvalues $\pm \frac{1}{2}$ with algebraic multiplicity 2. One can easily verify that for both $\frac{1}{2}$ and $-\frac{1}{2}$, the corresponding matrix $Z$ is indefinite. Then a random Hamiltonian perturbation with an arbitrary small 2-norm will move the eigenvalues off the imaginary axis. Note that

$$\|\mathcal{E}_2\|_2 = \frac{1}{2} + \mu < 1 = \|\mathcal{E}_1\|_2.$$

This example also demonstrates that the problem discussed in Example 3 of computing a minimum norm Hamiltonian perturbation that moves all purely imaginary eigenvalues of a Hamiltonian matrix from the imaginary axis is even more difficult. To achieve this it may be necessary to use a pseudo-spectral approach as suggested in [21].
Remark 2 So far we have studied complex Hamiltonian matrices and complex perturbations. For a real Hamiltonian matrix $H$ and real Hamiltonian perturbations we know that the eigenvalues occur in conjugate pairs. Then if $i\alpha (\alpha \neq 0)$ is an eigenvalue of $H$, so is $-i\alpha$, and from the real Hamiltonian Jordan form in Theorem 2.2 it follows that both eigenvalues have the same properties. Thus, in the real case one only needs to focus on the purely imaginary eigenvalues in the top half of the imaginary axis. It is not difficult to get essentially the same results as in Theorems 3.1, 3.2 for these purely imaginary eigenvalues.

The only case that needs to be studied separately is the eigenvalue 0. By Theorem 2.2 (a), if 0 is an eigenvalue of $H$, then it has either even-sized Jordan blocks, or pairs of odd-sized Jordan blocks with corresponding $Z$ indefinite, or both. This is the situation as in parts 2. and 3. of Theorem 3.1. So we conclude that 0 is a sensitive eigenvalue, meaning that a real Hamiltonian perturbation with arbitrary small norm will move the eigenvalue out of the origin, and there is no guarantee that the perturbed eigenvalues will be on the real or imaginary axis.

4 Perturbation of real eigenvalues of skew-Hamiltonian matrices

If $K$ is a complex skew-Hamiltonian matrix, then $iK$ is a complex Hamiltonian matrix. So the real eigenvalues of $K$ under skew-Hamiltonian perturbations behave in the same way as the purely imaginary eigenvalues of $iK$ under Hamiltonian perturbations. Theorems 3.1, 3.2 can be simply modified for $K$.

If $K$ is real and we consider real skew-Hamiltonian perturbations, then the situation is different. The real skew-Hamiltonian canonical form in Theorem 2.3 shows that each Jordan block occurs twice and the algebraic multiplicity of every eigenvalue is even. We obtain the following perturbation result.

**Theorem 4.1** Consider the skew-Hamiltonian matrix $K \in \mathbb{R}^{2n,2n}$ with a real eigenvalue $\alpha$ of algebraic multiplicity $2p$.

1. If $p = 1$, then for any skew-Hamiltonian matrix $E \in \mathbb{R}^{2n,2n}$ with sufficiently small $\|E\|$, $K+E$ has a real eigenvalue $\lambda$ close to $\alpha$ with algebraic and geometric multiplicity 2, which has the form

$$\lambda = \alpha + \eta + O(\|E\|^2),$$

where $\eta$ is the real double eigenvalue of the $2 \times 2$ matrix pencil $\lambda X^T J_n X - X^T (J_n E) X$, and $X$ is a full column rank matrix so that the columns of $X$ span the right eigenspace associated with $\alpha$. 

17
2. If there exists a Jordan block associated with $\alpha$ of size larger than 2, then generically for a given $\mathcal{E}$ some eigenvalues of $\mathcal{K} + \mathcal{E}$ will no longer be real.

If there exists a Jordan block associated with $\alpha$ of size 2, then for any $\epsilon > 0$, there always exists an $\mathcal{E}$ with $\|\mathcal{E}\| = \epsilon$ such that some eigenvalues of $\mathcal{K} + \mathcal{E}$ will be non-real.

3. If the algebraic and geometric multiplicity of $\alpha$ is equal and is greater than 2, then for any $\epsilon > 0$, there always exists an $\mathcal{E}$ with $\|\mathcal{E}\| = \epsilon$ such that some eigenvalues of $\mathcal{K} + \mathcal{E}$ will be non-real.

Proof. Based on the real skew-Hamiltonian canonical form in Proposition 2.3, we may determine a full column rank real matrix $X \in \mathbb{R}^{2n, 2p}$ such that span $X$ is the eigenspace corresponding to $\alpha$ and

$$X^T J_n X = J_p, \quad \mathcal{K} X = X \mathcal{R}, \quad \lambda(\mathcal{R}) = \{\alpha\}.$$ 

If $\|\mathcal{E}\|$ is sufficiently small, then for the perturbed matrix $\tilde{\mathcal{K}} = \mathcal{K} + \mathcal{E}$ and the associated $\tilde{X}$ and $\tilde{\mathcal{R}}$ such that $(\tilde{\mathcal{K}} + \mathcal{E}) \tilde{X} = \tilde{X} \tilde{\mathcal{R}}$, we have that

$$\|\tilde{X} - X\| = O(\|\mathcal{E}\|), \quad \tilde{X}^T J_n \tilde{X} = J_p,$$

where all the eigenvalues of $\tilde{\mathcal{R}}$ are close to $\alpha$. This implies that

$$\tilde{X}^T J_n (\mathcal{K} + \mathcal{E}) \tilde{X} = J_p \tilde{\mathcal{R}}, \quad (12)$$

and as in the Hamiltonian case we have the first order perturbation expression for $\tilde{\mathcal{R}}$ given by

$$\tilde{\mathcal{R}} = \mathcal{R} + J_p^T X^T (J_n \mathcal{E}) X + \delta E, \quad (13)$$

with $|\delta E| = O(\|\mathcal{E}\|^2)$.

We now prove part 1. When $p = 1$, then since the left hand side of (12) is real skew-symmetric, we have

$$\tilde{\mathcal{R}} = \lambda I_2,$$

where $\lambda$ is real. Clearly $\lambda$ is a real eigenvalue of $\mathcal{K} + \mathcal{E}$ and by assumption it is close to $\alpha$. So $\lambda$ has algebraic and geometric multiplicity 2 when $\|\mathcal{E}\|$ is sufficiently small.

Since $X^T (J_n \mathcal{E}) X$ is real skew-symmetric and $p = 1$ it follows that

$$J_p^T X^T (J_n \mathcal{E}) X = \eta I_2,$$

where $\eta$ is real. Obviously $\eta$ is an eigenvalue of $\lambda J_p - X^T (J_n \mathcal{E}) X$ with algebraic and geometric multiplicity 2. Note that the eigenvalues of $\lambda J_p - X^T (J_n \mathcal{E}) X$ are independent of the choice of $X$.

Using (13), parts 2. and 3. can be proved in the same way as the corresponding parts of Theorem 3.1.
5 Perturbation theory for the symplectic URV algorithm

In this section we will make use of the perturbation results obtained in Section 4 to analyze the perturbation of eigenvalues of real Hamiltonian matrices computed by the symplectic URV method proposed in [3].

For a real Hamiltonian matrix $H$, the symplectic URV method computes the factorization

$$U^T HV = R = \begin{bmatrix} R_1 & R_3 \\ 0 & R_2^T \end{bmatrix}, \quad (14)$$

where $U, V$ are real orthogonal symplectic, $R_1$ is upper triangular, and $R_2$ is quasi-upper triangular. Then, since $H$ is Hamiltonian, there exists another factorization

$$V^T HU = J R^T J = \begin{bmatrix} -R_2^T & R_3^T \\ 0 & -R_1^T \end{bmatrix}. \quad (15)$$

Combining (14) and (15), we see that

$$U^T H^2 U = \begin{bmatrix} -R_1 R_2^T & R_1 R_3^T - R_3 R_1^T \\ 0 & -(R_1 R_2^T)^T \end{bmatrix}.$$  

The matrix $H^2$ is real skew-Hamiltonian and its eigenvalues are the same as those of $-R_1 R_2^T$ but with double algebraic multiplicity. Note that the eigenvalues of $-R_1 R_2^T$ can be simply computed from its diagonal blocks. If $\gamma$ is an eigenvalue of $H^2$ then $\pm \sqrt{\gamma}$ are both eigenvalues of $H$.

In [3] the following backward error analysis was performed. Let $\hat{R}$ be the finite precision arithmetic result when computing $R$ in (14). Then there exist real orthogonal symplectic matrices $\hat{U}$ and $\hat{V}$ such that

$$\hat{U}^T (H + F) \hat{V} = \hat{R}, \quad (16)$$

where $F$ is a real matrix satisfying $|F| = c\|H\|\varepsilon$, with a constant $c$ and $\varepsilon$ the machine precision.

Similarly, there exists another factorization

$$\hat{V}^T (H + JF^T J) \hat{U} = J \hat{R}^T J.$$

and thus one has

$$\hat{U}^T (H^2 + E) \hat{U} = \hat{R} J \hat{R}^T J,$$

where

$$E = J^T [(JFH - (JFH)^T) - (JF)J(JF)^T]. \quad (17)$$

is real skew-Hamiltonian. So the computed eigenvalues that are determined from the diagonal blocks of $\hat{R} J \hat{R}^T J$ are the exact eigenvalues of $H^2 + E$, which is the real skew-Hamiltonian matrix $H^2$ perturbed by a small real skew-Hamiltonian perturbation $E$. We then obtain the following perturbation result.
Theorem 5.1 Suppose that $\lambda = i\alpha$ $(\alpha \neq 0)$ is a purely imaginary eigenvalue of a real Hamiltonian matrix $H$ and suppose that we compute an approximation $\hat{\lambda}$ by the URV method with backward errors $E$ and $F$ as in (17) and (16), respectively.

1. If $\lambda = i\alpha$ is simple, and $\|F\|$ is sufficiently small, then the URV method yields a computed eigenvalue $\hat{\lambda} = i\hat{\alpha}$ near $\lambda$, which is also simple and purely imaginary. Moreover, let $X$ be a real and full column rank matrix such that

\[HX = X \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}, \quad X^T J_n X = J_1,\]

and let

\[X^T J F X = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}.\]

Then

\[i\hat{\alpha} - i\alpha = -i \frac{f_{11} + f_{22}}{2} + O(\|F\|^2).\]

2. If $\lambda = i\alpha$ is not simple or $\lambda = 0$, then the corresponding computed eigenvalue may not be purely imaginary.

Proof. If $i\alpha$ is a nonzero purely imaginary eigenvalue of $H$, then so is $-i\alpha$. Thus $-\alpha^2$ is a negative eigenvalue of $H^2$ with algebraic and geometric multiplicity 2. Let $\hat{\mu}$ be the computed eigenvalue of $H^2$ that is close to $\mu = -\alpha^2$. Then $\hat{\mu}$ is also an eigenvalue of $H^2 + E$.

For part 1., if $\|F\|$ is sufficiently small, then by Theorem 4.1 part 1., $\hat{\mu}$ is real and negative. Thus we can express $\hat{\mu}$ as $\hat{\mu} = -\hat{\alpha}^2$ and $\hat{\alpha}$ has the same sign as $\alpha$. Note that

\[H^2 X = -\alpha^2 X,\]

and $\|E\| = O(\|F\|)$. Thus, by Theorem 4.1 we have

\[-\alpha^2 = -\alpha^2 + \eta + O(\|F\|^2),\]

where $\eta$ is the double real eigenvalue of

\[\lambda J_1 - X^T (J E) X.\]

Using (17) and that $HX = \alpha X J_1$, we obtain

\[X^T (J E) X = \alpha X^T J F X J_1 - \alpha J_1^T X^T F J^T J X + X^T F J F J^T J^T X = \alpha (X^T J F X J_1 - J_1^T X^T F J^T J X) + X^T F J F J^T J^T X = \alpha (f_{11} + f_{12}) J_1 + X^T F J F J^T J^T X,\]
which implies that
\[ \eta = \alpha (f_{11} + f_{12}) + O(\|F\|^2). \]

Then we have that
\[ -\hat{\alpha}^2 + \alpha^2 = \alpha (f_{11} + f_{22}) + O(\|F\|^2), \]
and thus
\[ \hat{\alpha} - \alpha = -\frac{\alpha}{\hat{\alpha} + \alpha} (f_{11} + f_{22}) + O(\|F\|^2), \]
which can be expressed as
\[ i\hat{\alpha} - i\alpha = -i \frac{f_{11} + f_{22}}{2} + O(\|F\|^2). \]

In part 2., if \( i\alpha \) is not simple and \( \alpha \neq 0 \), then \( -\alpha^2 \) is an eigenvalue of \( H^2 \) with algebraic multiplicity greater than 2. By parts 2. and 3. of Theorem 4.1 we cannot guarantee that the computed eigenvalue is still purely imaginary. If \( \alpha = 0 \) then by Remark 2, again there is no guarantee that the computed eigenvalues will be on the real axis or imaginary axis.

Note that the first order error bound of Theorem 5.1 was already given in [3] for simple eigenvalues of \( H \).

Although the symplectic URV method provides the computed spectrum of \( H \) with Hamiltonian symmetry, it can well approximate the location of the nonzero simple purely imaginary eigenvalues of \( H \). An analysis of the behavior of the method for multiple purely imaginary eigenvalues is an open problem.

Finally, we remark that the analysis can be easily extended to the method given in [4] for the complex Hamiltonian eigenvalue problem.

6 Conclusion

We have presented the perturbation analysis for purely imaginary eigenvalues of Hamiltonian matrices under Hamiltonian perturbations. We have shown that the structured perturbation theory yields substantially different results than the unstructured perturbation theory, in that it can (in some situations) be guaranteed that purely imaginary eigenvalues stay on the imaginary axis under structured perturbations while they generically move off the imaginary axis in the case of unstructured perturbations. These results show that the use of structure preserving methods can make a substantial difference in the numerical solution of some important practical problems arising in robust control, stabilization of gyroscopic systems as well as passivation of linear systems. We have also used the structured perturbation results to analyze the properties of the symplectic URV method of [3].
Acknowledgement

We thank R. Alam, S. Bora, D. Kressner, and J. Moro for fruitful discussions on structured perturbations, L. Rodman for providing further references, and L. Trefethen for initiating the discussion on the usefulness of structured perturbation analysis.

References


