Lipschitz stability of canonical Jordan bases of $H$-selfadjoint matrices under structure-preserving perturbations

T. Bella*, V. Olshevsky* and U. Prasad*

Abstract. In this paper we study Jordan-structure-preserving perturbations of matrices selfadjoint in the indefinite inner product. The main result of the paper is Lipschitz stability of the corresponding similitude matrices. The result can be reformulated as Lipschitz stability, under small perturbations, of canonical Jordan bases (i.e., eigenvectors and generalized eigenvectors enjoying a certain flipped orthonormality relation) of matrices selfadjoint in the indefinite inner product. The proof relies upon the analysis of small perturbations of invariant subspaces, where the size of a permutation of an invariant subspace is measured using the concepts of a gap and of a semigap.

1. Introduction. Part I. Preliminaries

1.1. Motivation and main result

Perturbation problems for matrices have been studied by many authors in different contexts, see, e.g., the monographs [GLR86, SS90, B97, KGMP03] among others, as well as the references therein. To motivate the problem considered in this paper we briefly recall several relevant results captured by the four cells (i) - (iv) of the following table.

Table 1. A selection of motivating results on the perturbation of Jordan structure.

<table>
<thead>
<tr>
<th></th>
<th>General perturbations (possibly changing Jordan structure)</th>
<th>Perturbations preserving Jordan structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>General matrices</td>
<td>(i)</td>
<td>(ii)</td>
</tr>
<tr>
<td>[GK78, MP80, DBT80, MO96]</td>
<td>[GR86, O89]</td>
<td></td>
</tr>
<tr>
<td>Matrices selfadjoint in indefinite inner product</td>
<td>(iii)</td>
<td>(iv)</td>
</tr>
<tr>
<td>[O91]</td>
<td>[GLR83, R06], this paper</td>
<td></td>
</tr>
</tbody>
</table>

Again, there is a vast literature on the subject, and the selection in the above table is clearly far from being comprehensive. It includes several references that directly motivate the problem considered.

(i). General matrices. General perturbations. It is well-known that even small perturbations of a given matrix $A_0$ can destroy its Jordan structure. For instance, for a nearby matrix $A$, not only the eigenvalues of $A$ but also the sizes of its corresponding Jordan blocks can be different from those of $A_0$. For a fixed $A_0$, the full description of all possible Jordan structures of nearby matrices $A$ was conjectured by Gohberg and Kaashoek in [GK78]. It was proven independently in [MP80, DBT80]. Two more proofs of the Gohberg-Kaashoek conjecture can be found in [MO96]. We do not discuss their general results in detail since in this paper we limit our focus to the special cases considered next.

(ii). General matrices. Perturbations preserving Jordan structure. In [GR86] (see also [GLR86]) the authors considered special perturbations $A$ that preserve the Jordan structure of $A_0$. We start with the following simplified version of their result.
Proposition 1.1 (Lipschitz stability of similarity matrices). Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed matrix. There is a constant $K > 0$ (depending on $A_0$ only) such that for any $A$ that is similar to $A_0$ there exists a similarity matrix $S$, i.e., $S^{-1}A_0S = A$, such that

$$\|I - S\| \leq K\|A - A_0\|.$$  \hfill (1.1)

In words, if a small perturbation $A$ of $A_0$ is similar to $A_0$, then a (highly nonunique) similarity matrix $S$ can be chosen to be a small perturbation of the identity matrix, and a Lipschitz–type bound (1.1) holds.

In fact, Gohberg and Rodman considered more general perturbations $A$ that are not similar to $A_0$ but have the same Jordan structure. Since the latter concept plays a key role in what follows we give two relevant definitions next.

Definition 1.2.

- (Same Jordan structure). Denote by $\sigma(A_0)$ and $\sigma(A)$ the sets of all eigenvalues of $A_0$ and $A$, respectively. Matrices $A_0$ and $A$ are said to have the same Jordan structure if there is a bijection $f : \sigma(A_0) \rightarrow \sigma(A)$ such that if $\mu = f(\lambda)$, then $\lambda$ and $\mu$ have the same Jordan block sizes.

- (Same Jordan bases). Matrices $A_0$ and $A$ that have the same Jordan structure are said to additionally have the same Jordan bases if the following statement is true. If $\mu = f(\lambda)$, then every Jordan chain of $A_0$ corresponding to $\lambda$ is also a Jordan chain of $A$ corresponding to $\mu$ (and automatically vice versa).

Remark 1.3 (Same Jordan bases). Two matrices $A_0, A \in \mathbb{C}^{n \times n}$ have the same Jordan bases if the following statement holds. If, for an invertible $T$, the matrix $T^{-1}A_0T$ is in a canonical Jordan form, then $T^{-1}AT$ is also in a canonical Jordan form.

In order to generalize Proposition 1.1 to perturbations $A$ having the same Jordan structure as $A_0$ we need to extend the concept of a similarity matrix $S$. The following obvious result is an enabling tool for doing this.

Lemma 1.4 (Similitude matrix). Two matrices $A_0, A \in \mathbb{C}^{n \times n}$ have the same Jordan structure if and only if there is an invertible matrix $S$ such that $A_0$ and $S^{-1}AS$ have the same Jordan bases.

We suggest to refer to the matrix $S$ in Lemma 1.4 as a similitude matrix since it generalizes the similarity matrix to the situation when $A_0$ and $A$ might not be similar but have the same Jordan structure. Observe that a similitude matrix is highly nonunique (just as its special case, a similarity matrix). We are now ready to present the following generalization of Proposition 1.1 that is implicit in [GR86, GLR86].

Proposition 1.5 (Lipschitz stability of similitude matrices). Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed matrix. There is a constant $K > 0$ (depending on $A_0$ only) such that for any $A$ having the same Jordan structure as $A_0$ there exists a similitude matrix $S$ (i.e., matrices $A_0$ and $S^{-1}AS$ have the same Jordan bases), such that

$$\|I - S\| \leq K\|A - A_0\|.$$  \hfill (1.2)

(iii). $H$-selfadjoint matrices. General perturbations. Matrices and their perturbations considered in the items (i) and (ii) above were general. It is of interest to study situations when both matrices $A_0$ and $A$ have some special structure. Hermitian structure is of little interest in the context of perturbations of Jordan structure since Hermitian (or selfadjoint) matrices are diagonalizable and they cannot have Jordan blocks. Matrices that are selfadjoint with respect to an indefinite inner product appear in a number of applications [GLR05], and they can have nontrivial Jordan blocks, so perturbation problems for their Jordan structure are of interest.

We refer to [GLR05] for a comprehensive introduction to the subject, and only recall here that for a Hermitian, invertible (not necessarily positive definite) matrix $H$, one defines the indefinite inner product by

$$[x, y]_H = (Hx, y) = y^*Hx, \quad \text{where } (x, y) = y^*x \text{ is the standard Euclidean inner product.}$$  \hfill (1.3)

Further, a matrix is $H$-selfadjoint (or selfadjoint with respect to $H$) if

$$[Ax, y]_H = [x, Ay]_H \quad (\text{for all } x, y \in \mathbb{C}^n), \quad \text{or, equivalently, } HA = A^*H.$$  \hfill (1.4)
We will use the notations \( \mathcal{S} \), \( \mathcal{A} \), \( \mathcal{H} \), \( \mathcal{K} \), \( \mathcal{D} \), \( \mathcal{M} \), \( \mathcal{O} \), \( \mathcal{G} \), and \( \mathcal{L} \) for the\( \mathcal{S} \)-selfadjoint (or Hermitian) matrices \( A \).

The monograph \([GLR83]\) contains a number of results on the perturbation of eigenvalues of \(H\)-selfadjoint matrices. The variation of the Jordan structure of \(H\)-selfadjoint matrices under small perturbations was studied in \([O91]\) where one can find certain restrictions additional to those of \([GK78, MP80, DBT80, MO96]\) mentioned in the item (i) above. The techniques used in \([O91]\) allow us to obtain an analog of Proposition 1.5 for \(H\)-selfadjoint matrices, which is described next.

(iv). **Main result.** **\(H\)-selfadjoint matrices. Perturbations preserving Jordan structure.** Let \(H_0\) be a fixed invertible Hermitian matrix, and let \(A_0\) be a fixed \(H_0\)-selfadjoint matrix. We consider their perturbations \(A\) and \(H\) where \(A\) is \(H\)-selfadjoint (in particular, \(H\) is invertible and Hermitian). This case was considered in \([GLR83, R06]\) where a number of results were obtained (we use some of them below). However, it seems the question of finding an analog of Proposition 1.5 has not been addressed in the literature yet. In order to obtain such an analog below one needs to carry over the concept of a similitude matrix to an indefinite inner product. The problem is that for an \(H\)-selfadjoint matrix \(A\), a similar matrix \(S^{-1}AS\) is not necessarily \(H\)-selfadjoint. This suggests that (in order to preserve the property of \(A\) of being selfadjoint with respect to indefinite inner product) the matrix \(H\) should also be modified appropriately. Here is the recipe.

**Definition 1.6 (Similarity-for-pairs relation).** Let \(A_0\) be \(H_0\)-selfadjoint and \(A\) be \(H\)-selfadjoint.

- We will use the notations
  \[
  (A, H) \xrightarrow{S} (A_0, H_0)
  \]
  to mean that \(S^{-1}AS = A_0\) and \(S^*HS = H_0\).

- The relation \(\langle (\cdot, \cdot) \xrightarrow{S} (\cdot, \cdot) \rangle\) will be called the similarity-for-pairs relation of matrices \((A, H)\), where \(A\) is \(H\)-selfadjoint.

Two remarks are due.

- A simple calculation shows that this notation makes sense; i.e., if \(A\) is \(H\)-selfadjoint and \(A, H) \xrightarrow{S} (B, G)\), then \(B\) is \(G\)-selfadjoint.

- It is easy to see that similarity-for-pairs is an equivalence relation.

In the above definition the matrices \(A_0\) and \(A\) are similar, so the corresponding \(S\) was indeed a similitude matrix. In the following definition we consider the case when \(A_0\) and \(A\) only have the same Jordan structure, and specify the concept of the similitude matrix for the indefinite inner product frameworks.

**Definition 1.7 (Weak similitude matrix).** Let \(A_0\) be \(H_0\)-selfadjoint and \(A\) be \(H\)-selfadjoint.

- A matrix \(S\) is called a (weak) similitude matrix of the quadruple \((A_0, H_0, A, H)\) if
  \[
  (A, H) \xrightarrow{S} (A_1, H_0),
  \]
  where matrices \(A_0\) and \(A_1\) have the same Jordan bases.

- In this case the pairs \((A_0, H_0)\) and \((A, H)\) are called (weakly) similitude.

With this background we can introduce the main result proved in the paper.

**Theorem 1.8 (Main result. Lipschitz stability of similitude matrices).** Let \(A_0 \in \mathbb{C}^{n \times n}\) be a fixed \(H_0\)-selfadjoint matrix. There exist constants \(K, \delta > 0\) (depending on \(A_0\) and \(H_0\) only) such that the following assertion holds. For any \(H\)-selfadjoint matrix \(A\) such that \(A\) has the same Jordan structure as \(A_0\) and
\[
\|A - A_0\| + \|H - H_0\| < \delta,
\]
the pairs \((A_0, H_0)\) and \((A, H)\) are similitude, and there exists a similitude matrix \(S\) such that
\[
\|I - S\| \leq K (\|A - A_0\| + \|H - H_0\|).
\]

In words, if (a) a small perturbation \(A\) of \(A_0\) has the same Jordan structure as \(A_0\); (b) \(H\) is a small perturbation of \(H_0\); (c) \(A_0\) is \(H_0\)-selfadjoint and \(A\) is \(H\)-selfadjoint, then (i) a similitude matrix \(S\) exists, and (ii) it can be chosen to be a small perturbation of the identity matrix, and a Lipschitz-type bound (1.7) holds.
Remark 1.9. There is a certain (deliberate) controversy here: while Definition 1.7 introduced a weak similitude matrix, the Theorem 1.8 asserts the existence of a (strong) similitude matrix to be formally introduced only in Definition 8.5 of Section 8. The controversy is only virtual since as we will see in Section 8, the weak similitude matrix \( S \) constructed in the course of proof of Theorem 1.8 will enjoy several additional nice properties. Hence the (strong) similitude matrix will be defined as a weak similitude matrix having those additional properties. For this reason, in Sections 2-7 the term “similitude” will be tentatively understood in the weak sense. In Section 8 it will be justified that all the results of the paper including Theorem 1.8 remain valid if the term “similitude” is understood in the strong sense. Therefore we will just use the nomenclature “similitude” without specifying whether it is weak or strong.

We preferred to formulate Theorem 1.8 before Definition 8.5 since the latter requires introducing a number of (unnecessary at the moment) technical details that will be dealt with in Sections 2 and 8.

Comparing Proposition 1.5 and Theorem 1.8, we see that the latter uses the assumption (1.6) not appearing in the former. We conclude this subsection with a simple example indicating that the condition (1.6) is essential, and it cannot simply be omitted.

Example 1.10 (Similitude matrix may not exist for large perturbations). Let us consider \( 1 \times 1 \) case:

\[
A_0 = \begin{bmatrix} 1 \end{bmatrix}, \quad H_0 = \begin{bmatrix} 1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 \end{bmatrix}, \quad H = \begin{bmatrix} -1 \end{bmatrix}.
\]

In this case the desired similitude matrix \( S \) does not exist, since it must satisfy \( S^* \cdot 1 \cdot S = -1 \). Clearly, \( 1 \times 1 \) matrices \( A_0 \) and \( A \) always have the same Jordan structure. However, \( H_0 \) and \( H \) are not close enough to ensure that (1.6) yields (1.7). In Section 2.1 we will recall another explanation [GLR05] of the fact that \( S \) does not exist here, it will be based on the concept of the so-called sign characteristic whose definition is recalled in Section 2.

1.2. Structure of the paper

The next section continues the introduction with three interpretations of Theorem 1.8, one of which is the second main result of the paper, Theorem 2.6. The section concludes with a graphical representation of the flow of the proofs of the paper. Section 3 presents a theorem showing it is sufficient for the proof of Theorem 1.8 to obtain the result for all pairs \((A_0, H_0)\) in the canonical form. In Section 4, Theorem 1.8 is proved in the case where \( A_0 \) consists of a single Jordan block corresponding to a real eigenvalue, or a pair of Jordan blocks corresponding to a single nonreal eigenvalue. Following this, Section 5 presents a decoupling result that allows the process of Section 4 to apply inductively. The proof of this result requires some auxiliary results on semigaps and gaps between subspaces which are given in Section 6. These results are then used to prove the results of Section 5 in Section 7. In Section 8, the second main result of the paper, Theorem 2.6 is proved, and the details of the distinction between weak and strong similitude introduced in Definition 1.7 and Remark 1.9 are explained in detail. Finally, in Section 9, the results of Theorems 1.8 and 2.6 are extended to the case of perturbations that partially preserve Jordan structure; that is, the sizes of Jordan blocks corresponding to some subset of the eigenvalues are unchanged.

2. Introduction. Part II. Three interpretations of Theorem 1.8

The second part of the introduction is somewhat more technical. In Sections 2.2 and 2.3 below we provide three useful interpretations of our main result. They will use two key concepts defined next.

2.1. Key definitions. Sign characteristic and canonical Jordan bases

We begin with quoting a fundamental theorem of [W68, M63, GLR05] which plays a central role in all arguments below. As usual, \( J(\lambda) \) denotes a single Jordan block of the form

\[
J(\lambda) = \begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 1 \\
0 & \cdots & 0 & \cdots & \lambda
\end{bmatrix}, \quad \tilde{J}(\lambda) = \begin{cases}
J(\lambda) & \lambda \in \mathbb{R} \\
\begin{bmatrix}
J(\lambda) & 0 \\
0 & J(\lambda)
\end{bmatrix} & \lambda \notin \mathbb{R}
\end{cases}, \quad \tilde{I} = \begin{bmatrix}
0 & \cdots & 0 & 1 \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 1 & 0 \\
1 & \cdots & \cdots & \cdots
\end{bmatrix}
\]

(2.1)

The matrix \( \tilde{I} \) was called the sip (standard involuntary permutation) matrix in [GLR86].
Theorem 2.1 ([GLR05], Theorem 5.1.1). (Canonical form of matrices selfadjoint with respect to indefinite inner product). Let $A \in \mathbb{C}^{n \times n}$ be a fixed $H$-selfadjoint matrix for some invertible, selfadjoint matrix $H$. Then there exists an invertible matrix $T$ such that

$$\begin{align*}
(A, H) & \overset{T}{\rightarrow} (J, P) \\
\end{align*}$$

(2.2)

where

$$J = J(\lambda_1) \oplus \cdots \oplus J(\lambda_n) \oplus \tilde{J}(\lambda_{\alpha+1}) \oplus \cdots \oplus \tilde{J}(\lambda_{\beta})$$

(2.3)
is a Jordan normal form of $A$ for real eigenvalues $\lambda_1, \ldots, \lambda_n$ and nonreal eigenvalues $\lambda_{\alpha+1}, \ldots, \lambda_{\beta}$ from the upper half-plane, and

$$P = P_1 \oplus \cdots \oplus P_{\alpha} \oplus P_{\alpha+1} \oplus \cdots \oplus P_{\beta}$$

(2.4)

where $P_k$ is a signed sip matrix $\epsilon_k \tilde{I}$ of the same size as $J(\lambda_k)$ (for $k = 1, \ldots, \alpha$), and a sip matrix $\tilde{I}$ of the same size as $\tilde{J}(\lambda_k)$ (for $k = \alpha + 1, \ldots, \beta$), and $\epsilon_k = \pm 1$ for $k = 1, \ldots, \alpha$. The set

$$\epsilon = \{\epsilon_1, \ldots, \epsilon_{\alpha}\}$$

(2.5)
is determined by the pair $(A, H)$ uniquely, up to a permutation of the signs $\epsilon_k$ corresponding to Jordan blocks of the same size and of the same eigenvalue.

Remark 2.2 (Symmetry of eigenvalues). In particular, nonreal eigenvalues of an $H$-selfadjoint matrix $A$ come in complex conjugate pairs. Furthermore, for each nonreal conjugate pair of eigenvalues the sizes of their Jordan blocks are identical.

Definition 2.3 (Canonical form and sign characteristic, [GLR83]). The pair $(J, P)$ in (2.3) and (2.4) is called a canonical form of $(A, H)$. The set of signs in (2.5) is called the sign characteristic of the pair $(A, H)$.

Recall that “similarity for pairs” is an equivalence relation, and hence pairs that have different canonical forms (up to an appropriate rearrangement of Jordan blocks of $A$ and corresponding blocks of $H$) cannot be similar. This is exactly what happened in Example 1.10. Indeed, it is immediate to see that the pairs $(A_0, H_0)$ and $(A, H)$ are in the canonical form, from which we can see they have different sign characteristics, and therefore they can not be similar.

The first equation (2.2) implies $T^{-1}AT = J$, which means that the columns of the matrix $T$ form a Jordan basis of $A$. However, not all such matrices $T$ satisfy the second equation $T^{*}HT = P$, also implied by (2.2), with $P$ of (2.4). We coin a special name for the columns of those matrices $T$ that satisfy both equations implied by (2.2).

Definition 2.4 (Canonical Jordan basis of an $H$-selfadjoint matrix). Let $A$ be an $H$-selfadjoint matrix, and let $T$ be a similarity matrix that brings $(A, H)$ to its canonical form $(J, P)$. The columns of $T$ form a Jordan basis of $A$ that will be called a canonical Jordan basis of $(A, H)$.

The following example makes the property of “flipped orthonormality” of the vectors of a canonical basis more transparent.

Example 2.5 (Flipped orthonormality). Let

$$J = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad P = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

be a canonical pair. The canonical Jordan basis \{\{e_1, e_2, e_3\}, \{e_4, e_5\}\} of $(J, P)$ consists of two Jordan chains

$$0 \leftarrow e_1 \leftarrow e_2 \leftarrow e_3, \quad 0 \leftarrow e_4 \leftarrow e_5.$$  

It is easy to see that the canonical structure of $P$ yields that vectors belonging to different chains of this canonical basis are $P$-orthogonal, i.e.,

$$[e_j, e_k]_P = 0, \quad (j = 1, 2, 3; k = 4, 5).$$

Further, for the same reason the vectors within one chain have what might be called “flipped orthonormality”:

$$[e_j, e_k]_P = \delta_{j,k} - \delta_{j,4} \quad (j, k = 1, 2, 3), \quad [e_j, e_k]_P = \delta_{j,3} - \delta_{j,5} \quad (j, k = 4, 5).$$

It is the above “flipped orthonormality property” that distinguishes canonical Jordan basis from the other ones.
With this background we are now ready to present a first interpretation of Theorem 1.8.

2.2. First interpretation of the main result. Lipschitz stability of canonical Jordan bases of $H$-selfadjoint matrices under small perturbations preserving Jordan structure

Let $\{\lambda_1, \ldots, \lambda_3\}$ be a set of all eigenvalues of $A_0$. Denote

$$m_k(A_0, \lambda_s) := \text{the length of the } k\text{-th Jordan chain } \{f^{(k,s)}_r \}_{r=0}^{m_k(A_0, \lambda_s) - 1}\$$

of the matrix $A_0$ corresponding to its eigenvalue $\lambda_s$. Throughout the paper we assume that $\{m_k(A_0, \lambda_s)\}$ are ordered in nonascending order. Let $A$ have the same Jordan structure as $A_0$, which means that the eigenvalues $\{\mu_1, \ldots, \mu_3\}$ of $A$ can be ordered such that

$$m_k(A_0, \lambda_s) = m_k(A, \mu_s).$$

With these notations Theorem 1.8 implies the following result on stability of eigenvectors and generalized eigenvectors.

Theorem 2.6 (Lipschitz stability of canonical Jordan bases). Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed $H_0$-selfadjoint matrix. Let

$$\{\{f^{(k,s)}_r \}_{r=0}^{m_k(A_0, \lambda_s) - 1} \}_{s=\beta, k=\text{dim Ker}(A_0 - \lambda_s I)}$$

be a fixed canonical Jordan basis of $A_0$. There exist constants $K, \delta > 0$ (depending on $A_0$ and $H_0$ only) such that the following assertion holds. For any $H$-selfadjoint matrix $A$ such that $A$ has the same Jordan structure as $A_0$ and

$$\|A - A_0\| + \|H - H_0\| < \delta,$$

there exists a canonical Jordan basis

$$\{\{g^{(k,s)}_r \}_{r=0}^{m_k(A, \lambda_s) - 1} \}_{s=\beta, k=\text{dim Ker}(A - \lambda_s I)}$$

of $A$ such that

$$\|g^{(k,s)}_r - f^{(k,s)}_r\| \leq K \left(\|A - A_0\| + \|H - H_0\|\right)$$

for all $k, s, r$ within their ranges.

In words, let $\{f^{(k,s)}_r \}$ be a fixed canonical Jordan basis of a fixed $H_0$-selfadjoint matrix $A_0$, and let $(A, H)$ be a small perturbation of $(A_0, H_0)$ where $A$ is $H$-selfadjoint. If $A$ has the same Jordan structure as $A_0$, then a (highly nonunique) canonical Jordan basis $\{g^{(k,s)}_r \}$ of $(A, H)$ can be chosen to be a small perturbation of the given canonical Jordan basis of $(A_0, H_0)$, and the Lipschitz-type bound (2.8) holds.

The proof of the above result will be given later in Section 8.1.

2.3. Second interpretation of the main result. Similitude matrix is $(H_0, H)$-unitary for small perturbations

Let us specify Theorem 1.8 to the case when $H_0 = H = I$. It is easy to see from the definition in (1.4) that in this case both matrices $A_0$ and $A$ are Hermitian. Secondly, in this case (1.5) implies $S^*S = I$, i.e., $S$ is a unitary matrix. To sum up, Theorem 1.8 specifies to the following result.

Theorem 2.7. Let $A_0$ be a fixed Hermitian matrix. Then there exists a constant $K > 0$ such that for any Hermitian matrix $A$ there exists a unitary similitude matrix $S$ (i.e., such that $A_0$ and $S^{-1}AS$ are simultaneously diagonalizable) such that

$$\|I - S\| \leq K \|A - A_0\|.$$

In words, if $A_0$ and $A$ are both Hermitian, then the similitude matrix $S$ can be chosen to be unitary and satisfying the bound (2.9). The latter result (Lipschitz stability of eigenvectors of Hermitian matrices) is known (e.g., it is an obvious consequence of [RP87]), but it leads to an interesting interpretation (cf. with [R06]) of the main result of the paper, Theorem 1.8.

In this context, the meaning of Theorem 1.8 is that extending to the case of indefinite inner products, under the stated conditions the similitude matrix $S$ can be chosen to be $(H_0, H)$-unitary (i.e., $S^*HS = H_0$) and satisfying the Lipschitz-type bound (1.7).
2.4. Third interpretation of the main result. Lipschitz stability of congruency matrices

In Section 2.3 we considered a special case when the matrix \( H \) in \((A, H)\) was the identity matrix, i.e., \( H_0 = H = I \). Here we consider another special case and set \( A_0 = A = I \). Clearly, \( I \) is \( H \)-selfadjoint for any invertible Hermitian \( H \). Here is a specialization of our main result, Theorem 1.8, in this case.

**Theorem 2.8 (Lipschitz stability of congruency matrices).** Let \( H_0 \) be a fixed invertible Hermitian matrix. There exist constants \( K, \delta > 0 \) (depending on \( H_0 \) only) such that the following assertion holds. For any Hermitian matrix \( H \) such that

\[
\| H - H_0 \| < \delta,
\]

there exists a congruency matrix \( S \), i.e., \( S^* H S = H_0 \), such that

\[
\| I - S \| \le K \| H - H_0 \|.
\]

Note that in view of (2.10) and (2.11) both \( H \) and \( S \) are invertible, and \( \| S \| \) and hence \( \| S^{-1} \| \) are bounded. Hence

\[
\| I - S^{-1} \| = \| S^{-1} (S - I) \| \le \| S^{-1} \| \cdot \| I - S \| = K_1 \| H - H_0 \|.
\]

with some \( K_1 \). This is a local version of the main result of [RP87] who proved the bound \( \| I - S^{-1} \| \le K_1 \| H - H_0 \| \) without the restriction\(^1\) (2.10), but requiring instead that \( H_0 \) and \( H \) remain congruent.

2.5. The flow of the results

The main results of the paper are Theorems 1.8 and 2.6. The following diagram presents the flow of the proofs.

---

![Diagram](image)

---

\(^1\)Note that examples indicate that without restriction (2.10) the bound (2.11) does not hold.
3. Reduction to the canonical form

We will find it useful throughout the paper to assume that the matrices \( A_0 \) and \( H_0 \) of Theorem 1.8 are in the canonical form. There is no loss of generality with this assumption, as the next theorem demonstrates.

**Theorem 3.1 (Reduction to the canonical form).** Suppose the result of Theorem 1.8 is true for each pair \( (A_0, H_0) \) in the canonical form as defined in the Definition 2.3. Then the result of Theorem 1.8 is true for all pairs \( (B_0, G_0) \), where \( B_0 \) is \( G \)-selfadjoint.

**Proof.** Suppose the pair \( (B_0, G_0) \) is not in the canonical form. By Theorem 2.1, there exists a matrix \( T \) such that

\[
(B_0, G_0) \rightarrow (A_0, H_0)
\]

for some pair \( (A_0, H_0) \) in the canonical form. Then we have

\[
(A, H) \xrightarrow{\sigma} (A_1, H_0)
\]

\[
(T) \rightarrow T
\]

\[
(B, G) \xrightarrow{R=TST^{-1}} (B_1, G_0)
\]

The above diagram implies that the bound for a general \( (B_0, G_0) \),

\[
\|I - R\| \leq K_{B_0, G_0} (\|B - B_0\| + \|G - G_0\|),
\]

(3.2)

can be deduced from the bound for a canonical pair \( (A_0, H_0) \),

\[
\|I - S\| \leq K_{A_0, H_0} (\|A - A_0\| + \|H - H_0\|).
\]

Indeed, using the standard notation \( \kappa(T) = \|T\| \cdot \|T^{-1}\| \) and the formulas captured by the diagram (3.1) we compute

\[
\|I - R\| = \|I - TST^{-1}\| \leq \kappa(T) \|I - S\| \leq \kappa(T) K_{A_0, H_0} (\|A - A_0\| + \|H - H_0\|)
\]

\[
\leq \kappa(T) K_{A_0, H_0} (\|T^{-1}BT - T^{-1}B_0T\| + \|T^*GT - T^*G_0T\|)
\]

\[
= \kappa(T) K_{A_0, H_0} (\|B - B_0\| + \|T\| \|G - G_0\|)
\]

\[
\leq K_{B_0, G_0} (\|B - B_0\| + \|G - G_0\|).
\]

Hence it suffices to consider in what follows only the cases where \( (A_0, H_0) \) are in the canonical form.

4. Perturbations of a single real Jordan block or of a pair of complex conjugate Jordan blocks

In this section we present the first step of the proof of Theorem 1.8 for the special case where

\[
A_0 = \tilde{J}(\lambda), \quad H_0 = \epsilon \tilde{I}, \quad \text{(with } \epsilon = \pm 1),
\]

(4.1)

where the sip matrix \( \tilde{I} \) and Jordan block \( \tilde{J}(\lambda) \) were defined in (2.1). Recall that for a real \( \lambda \) the matrix \( \tilde{J}(\lambda) \) is a single Jordan block, and for a nonreal \( \lambda \) we have \( \epsilon = 1 \) and the matrix \( \tilde{J}(\lambda) \) is a direct sum of two Jordan blocks. In both cases it is easy to see that \( A_0 \) is \( H_0 \)-selfadjoint.

**Theorem 4.1 (Lipschitz stability of similitude matrices in the single block case).** Let \( A_0 \in \mathbb{C}^{n \times n} \) be a fixed \( H_0 \)-selfadjoint matrix as given in (4.1). There exist constants \( K, \delta > 0 \) (depending on \( A_0 \) and \( H_0 \) only) such that the following assertion holds. For any \( H \)-selfadjoint matrix \( A \) such that \( A \) has the same Jordan structure as \( A_0 \) and

\[
\|A - A_0\| + \|H - H_0\| < \delta,
\]

(4.2)

there exists a similitude matrix \( S \) such that

\[
\|I - S\| \leq K (\|A - A_0\| + \|H - H_0\|).
\]

(4.3)

Theorem 4.1 will be proved in Section 4.2, and an extension of it to the case of a single complex eigenvalue will be discussed in Section 4.3. Before doing so, we illustrate three of the steps by which the proof will proceed with the following simple example.
4.1. Single–Jordan–block model example (for Theorem 4.1)

We begin this section with a simple example involving a matrix $A_0$ in Jordan form. For the pair

$$A_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad H_0 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

consider the matrices

$$A = \begin{bmatrix} 0 & 1 & 2\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2\delta \\ 1 & -2\delta & \delta \end{bmatrix}$$

for some small positive $\delta$. The pair $(A, H)$ is then a small perturbation of the pair $(A_0, H_0)$, and it is straightforward to verify that $A_0$ is $H_0$-selfadjoint and $A$ is $H$-selfadjoint. We wish to produce a matrix $S$ satisfying

(i) $(A, H) \overset{S}{\longrightarrow} (A_0, H_0)$

(ii) $\|I - S\| \leq K(\|A - A_0\| + \|H - H_0\|)$.

We will design this matrix $S$ in three steps, the first will produce a matrix $S_1$, the second will produce a matrix $S_2$, and the third step will combine these as $S = S_1S_2$, and check that it satisfies the desired bound.

4.1.1. First step. Mapping $A \rightarrow A_1$. Constructing $S_1$. Notice that the Jordan chain of $A$ corresponding to $\lambda = 0$ is

$$0 \leftarrow e_1 \leftarrow e_2 + (2\delta)e_1 \leftarrow e_3,$$

where $\leftarrow$ denotes application of the matrix $A$. This chain is a small perturbation of that of $A_0$ corresponding to $\lambda = 0$, which is simply

$$0 \leftarrow e_1 \leftarrow e_3 \leftarrow e_3.$$

The matrix that maps these basis vectors (those of $A_0$ into those of $A$) is given by

$$S_1 = \begin{bmatrix} 1 & 2\delta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we have

$$S_1^{-1}AS_1 = A_0, \quad \text{and} \quad S_1^*HS_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} =: H_1$$

so that

$$(A, H) \overset{S_1}{\longrightarrow} (A_0, H_1).$$

This illustrates the need for the second step below, as we need to generate an $S$ such that $(A, H) \overset{S}{\longrightarrow} (A_0, H_0)$ and $H_0 \neq H_1$.

4.1.2. Second step. Zeroing sub–antidiagonal entries of $H_1$. Constructing $S_2$. The next step is to choose a matrix $S_2$ so that $(A_0, H_1) \overset{S_2}{\longrightarrow} (A_0, H_0)$, that is, a matrix that repairs the problem below the anti-diagonal in $H_1$ of (4.4) to produce $H_0$ without modifying the fact that $S_1^{-1}AS_1$ already produced $A_0$. The existence of such a matrix $S_2$ in general will be proven in the coming sections, but for now, notice that the matrix

$$S_2 = \begin{bmatrix} 1 & 0 & -\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

satisfies the required conditions; that is

$$S_2^*H_1S_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = H_0, \quad \text{and} \quad S_2^{-1}A_0S_2 = S_2^{-1}S_2A_0 = A_0.$$
so that

\[(A_0, H_1) \xrightarrow{S_2} (A_0, H_0).\]

### 4.1.3. Third step. Combining \(S_1\) and \(S_2\).
Set \(S = S_1S_2\). We have demonstrated in the previous two steps that

\[(A, H) \xrightarrow{S_1} (A_0, H_1) \xrightarrow{S_2} (A_0, H_0) \quad \text{and hence} \quad (A, H) \xrightarrow{S} (A_0, H_0),\]

which implies condition (i). Furthermore, computing \(S\) explicitly yields

\[S = S_1S_2 = \begin{bmatrix} 1 & 2\delta & -\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.\]

Hence all three differences \(A - A_0, H - H_0,\) and \(I - S\) are of the same small order of \(\delta\), and so condition (ii) is satisfied.

### 4.2. Proof of Theorem 4.1 in the case of a real single Jordan block.
In this section we illustrate that the approach demonstrated in the above example can yield the proof of Theorem 4.1. The presentation follows that of the example in the previous section and is organized into three sections.

#### 4.2.1. First step. Mapping \(A \rightarrow A_1\). Constructing \(\tilde{S}_1\).
In this section we prove the following proposition, which asserts the existence of a matrix \(\tilde{S}_1\) as in the first step in the above example earlier. (Note: The tilde indicates that this is not exactly the matrix \(S_1\) mentioned in the example above, but this difference is explained and handled in Section 4.2.2.)

**Proposition 4.2 (Constructing \(\tilde{S}_1\)).** Let \(A_0 \in \mathbb{C}^{n \times n}\) be a fixed \(H_0\)-selfadjoint matrix as given in (4.1). There exists a constant \(K > 0\) (depending on \(A_0\) and \(H_0\) only) such that the following assertion holds. For any \(H\)-selfadjoint matrix \(A\) that is similar to \(A_1 = J_n(\mu)\) for some \(\mu\) there exists a similitude matrix \(\tilde{S}_1\) such that

\[(A, H) \xrightarrow{\tilde{S}_1} (A_1, \tilde{H}_1) \quad \text{for some lower anti–triangular Hankel matrix}\ \tilde{H}_1 \ \text{of the form}\]

\[\begin{bmatrix} 0 & \cdots & 0 & \ast \\ \vdots & \ddots & \ddots & \ast \\ 0 & \ast & \ddots & \vdots \\ \ast & \ast & \cdots & \ast \end{bmatrix}\]

satisfying

\[\|I - \tilde{S}_1\| \leq K\|A - A_0\|\].

The proposition will be proved by the following lemma.

**Lemma 4.3 (Bounds for the perturbed Jordan chains).** Let \(A_0 \in \mathbb{C}^{n \times n}\) have only one eigenvalue \(\lambda\). Let \(\{f_k\}_{k=0}^{m-1}\) be the longest Jordan chain corresponding to the eigenvalue \(\lambda\) of \(A_0\). There are constants \(K, \delta > 0\) (depending on \(A_0\) only) such that for any \(A \in \mathbb{C}^{n \times n}\) having only one eigenvalue \(\mu\) and satisfying

\[\|A - A_0\| \leq \delta\]

we have the following. If the maximal length of the corresponding Jordan chain of \(A\) is also \(m\), then there exists a Jordan chain \(\{g_k\}_{k=0}^{m-1}\) of \(A\) such that

\[\|f_k - g_k\| \leq K\|A - A_0\|, \quad k = 0, \ldots, m - 1.\]

This lemma is actually a specification of its more general version, Lemma 6.14, which will be proved later without using any intermediate results.

We are now ready to provide the proof of Proposition 4.2.

---

3By lower anti-triangular Hankel matrix we mean a matrix whose entries above the anti-diagonal are zeros.
Proof of Proposition 4.2. Using the notations of Lemma 4.3, choose the matrix $\tilde{S}_1$ such that $\tilde{S}_1 f_i = g_i$, for $i = 0, \ldots, n - 1$. We denote the result of applying the matrix $\tilde{S}_1$ to $H$ by $\tilde{H}_1$, so

$$(A, H) \xrightarrow{\tilde{S}_1} (A_1, \tilde{H}_1).$$

We next show that $\tilde{H}_1$ is lower anti–triangular Hankel. Writing $A_1 = \mu I + Z^T$ where $Z$ denotes the lower shift matrix, and from the fact that $A_1$ is $\tilde{H}_1$-selfadjoint we have that

$$\tilde{H}_1(\mu I + Z^T) = (\mu I + Z)\tilde{H}_1$$

and hence

$$\tilde{H}_1 Z^T = Z \tilde{H}_1,$$

from which it immediately follows that $\tilde{H}_1$ is lower anti–triangular Hankel.

Next, for any $x \in \mathbb{C}^n$ with $\|x\| = 1$, write $x = \sum_{i=0}^{n-1} \alpha_i f_i$. Then $y = \tilde{S}_1 x = \sum_{i=0}^{n-1} \alpha_i g_i$, and

$$\| (I - \tilde{S}_1) x \| = \| x - y \| \leq \sum_{i=0}^{n-1} |\alpha_i| \| f_i - g_i \| \leq \max_{0 \leq i \leq n-1} |\alpha_i| \cdot \| f_i - g_i \|.$$

Next, denoting by $[x]_F = (\alpha_{-1})$ the coordinates of $x$ with respect to the fixed basis $\{f_i\}$ and $F_{-E}$ the change of basis matrix from the standard basis $\{e_i\}$ to $\{f_i\}$, we have

$$\max_{0 \leq i \leq n-1} |\alpha_i| \leq \sqrt{\sum_{j=0}^{n-1} |\alpha_j|^2} = \| [x]_F \| \leq \| F_{-E} \| \cdot \| x \| = \| F_{-E} \|. $$

Using Lemma 4.3 we have that

$$\| (I - \tilde{S}_1) x \| \leq \| F_{-E} \| \cdot \| f_i - g_i \| \leq K \| A - A_0 \|$$

for any $x \in \mathbb{C}^n$ with $\|x\| = 1$. Hence

$$\| I - \tilde{S}_1 \| \leq K \| A - A_0 \| \tag{4.5}$$

as desired. $\square$

4.2.2. Modified first step. Forcing unit antidiagonal of $H_1$. Constructing $S_1$. The matrix $\tilde{S}_1$ constructed in Section 4.2.1 was such that $(A, H) \xrightarrow{\tilde{S}_1} (A_1, \tilde{H}_1)$ with $\tilde{H}_1$ a lower anti–triangular Hankel matrix of the form

$$\tilde{H}_1 = \epsilon \cdot \begin{bmatrix} 0 & \cdots & 0 & a \\ \vdots & \ddots & a & \ast \\ 0 & \ddots & \ddots & \vdots \\ a & \ast & \ddots & \ast \end{bmatrix} \quad \text{with} \quad \epsilon = \pm 1. \tag{4.6}$$

In the next proposition we construct a different matrix $S_1$ such that mapping $(A, H) \xrightarrow{S_1} (A_1, H_1)$ produces a better matrix $H_1$ of the form

$$H_1 = \epsilon \cdot \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & 1 & \ast \\ 0 & \ddots & \ddots & \vdots \\ 1 & \ast & \ddots & \ast \end{bmatrix} \quad \text{with} \quad \epsilon = \pm 1. \tag{4.7}$$

(Recall that in the model example of this section the matrix $H_1$ of (4.4) indeed had the form (4.7)).

Proposition 4.4 (Modified Proposition 4.2. Constructing $S_1$). Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed $H_0$-selfadjoint matrix as given in (4.1). There exist constants $K, \delta > 0$ (depending on $A_0$ and $H_0$ only) such that the following assertion holds. For any $H$-selfadjoint matrix $A$ that is similar to $A_1 = J_n(\mu)$ for some $\mu$ and

$$\| A - A_0 \| + \| H - H_0 \| < \delta,$$

there exists a similitude matrix $S_1$ such that

$$(A, H) \xrightarrow{S_1} (A_1, H_1) \quad \text{for some lower anti–triangular Hankel matrix } H_1 \text{ of the form (4.7)}$$
satisfying

\[ \|I - S_1\| \leq K (\|A - A_0\| + \|H - H_0\|). \]

Let \( \tilde{S}_1 \) be the matrix guaranteed by Proposition 4.2, and \( a \) the antidiagonal entry of the resulting matrix \( \tilde{H}_1 \) as in (4.6). We will demonstrate that the matrix

\[ S_1 := a^{-1/2} \tilde{S}_1 \]

is the desired matrix. Indeed, it is easy to see that \( S_1^{-1} AS_1 = A_1 \) and \( S_1^* HS_1 = H_1 \) for \( H_1 \) of the form (4.7).

The next lemma demonstrates that the bound on \( \|I - S_1\| \) given by Proposition 4.2 implies the corresponding bound for \( \|I - S_1\| \).

**Lemma 4.5 (Forcing unit antidiagonal of \( G \)).** Let \( A_0 \in \mathbb{C}^{n \times n} \) be an \( H_0 \)-selfadjoint matrix. There exist positive constants \( K_1, K_2, \delta > 0 \) (all depending on \( A_0 \) and \( H_0 \) only) such that the following assertion holds. For any \( H \)-selfadjoint matrix \( A \) satisfying

\[ \|A - A_0\| + \|H - H_0\| \leq \delta, \tag{4.8} \]

and any invertible matrix \( T \) satisfying

\[ \|I - T\| \leq K_1 (\|A - A_0\| + \|H - H_0\|), \tag{4.9} \]

and such that the matrix \( B \) in \( (A, H) \xrightarrow{T} (B, G) \) is a single Jordan block we have

\[ \|I - a^{-1/2} T\| \leq K_2 (\|A - A_0\| + \|H - H_0\|), \tag{4.10} \]

where \( a \) is the \((n, 1)\)-entry of the matrix \( G \).

**Proof.** The proof consists of two parts. In part (i) we prove that there is \( K_3 > 0 \) (depending on \( A_0 \) and \( H_0 \) only) such that

\[ \|G - H_0\| \leq K_3 (\|A - A_0\| + \|H - H_0\|) \quad \text{with} \quad K_3 = ((1 + K_1 \delta)^2 + (2 + K_1 \delta)\|H_0\|). \tag{4.11} \]

Then we prove in part (ii) that the desired bound (4.10) is valid.

(i). **(Proving (4.11))** We have

\[ \|G - H_0\| = \|T^* HT - H_0\| = \|T^* HT - T^* H_0 T + T^* H_0 T - T^* H_0 + T^* H_0 - H_0\| \leq \|T\|^2 \cdot \|H - H_0\| + (1 + \|T\|)(\|H_0\| \cdot \|T - I\|). \tag{4.12} \]

From conditions (4.8) and (4.9), we have

\[ \|T\| \leq 1 + K_1 (\|A - A_0\| + \|H - H_0\|) \leq 1 + K_1 \delta =: M \tag{4.13} \]

and substituting this estimate in (4.12) yields

\[ \|G - H_0\| \leq [M^2 + (1 + M)\|H_0\|] (\|A - A_0\| + \|H - H_0\|), \]

which yields (4.11).

(ii). **(Proving (4.10))** We first show that

\[ |a - 1| \leq \tilde{K}_2 (\|A - A_0\| + \|H - H_0\|), \quad a > \frac{1}{2}, \tag{4.14} \]

for some constant \( \tilde{K}_2 \). Using the relations

\[ \epsilon = e_n^T H_0 e_1 \quad \text{and} \quad \epsilon \cdot a = e_n^T G e_1 \]

with \( e_k \) the \( k \)-th unit vector, and the fact that \( \epsilon = \pm 1 \), we have that

\[ |a - 1| = |\epsilon \cdot a - \epsilon| = |e_n^T (G - H_0) e_1| \leq \|G - H_0\| \]

and the bound for \( |a - 1| \) follows from (4.11). Next, since \( \delta \) is at our disposal, we can always assume that it is small enough to guarantee that

\[ \delta \leq (2\tilde{K}_2)^{-1}, \]

so the condition (4.8) implies that

\[ |a - 1| \leq \tilde{K}_2 (\|A - A_0\| + \|H - H_0\|) < \tilde{K}_2 \delta \leq \frac{1}{2}, \]

and hence \( a > 1/2 \), so both statements of (4.14) follow by perhaps considering a smaller neighborhood \( \delta \) of the pair \((A_0, H_0)\).
Proposition 4.6 (Zeroing sub–antidiagonal entries of
proposition implements the second step of the model example.

That is, it must not affect the matrix \( A \), and it must zero out the sub–antidiagonal elements of the matrix \( H_1 \). The former can be accomplished by choosing \( S_2 \) to be upper triangular Toeplitz, as since \( A_1 \) is also upper triangular Toeplitz, they will commute. For the latter, this is essentially accomplished by choosing a matrix that is the matrix square root of the inverse of the given Hankel matrix \( H_1 \). The next lemma shows that this matrix has the desired properties.

Next,

\[
I - \frac{1}{\sqrt{a}} T = I - T + T - \frac{1}{\sqrt{a}} T \leq \| I - T \| + \left| \frac{\sqrt{a} - 1}{\sqrt{a}} \right| \| T \|,
\]

and using (4.14), we have that

\[
\left| \frac{\sqrt{a} - 1}{\sqrt{a}} \right| = \left| \frac{a - 1}{a + \sqrt{a}} \right| \leq \frac{a - 1}{a} \leq 2|a - 1|.
\]

Thus, we have that

\[
\left\| I - \frac{1}{\sqrt{a}} T \right\| \leq \| I - T \| + 2|a - 1| \cdot \| T \|
\]

and using the bounds for \( \| T \| \) in (4.13) and \( |a - 1| \) in (4.14), we arrive at

\[
\| I - a^{-1/2} T \| \leq [K_1 + 2M(M^2 + (1 + M))\|H_0\|] (\|A - A_0\| + \|H - H_0\|)
\]

which establishes desired (4.10). \( \square \)

We next return to and prove Proposition 4.4.

Proof of Proposition 4.4. For the given pair \((A_0, H_0)\), Proposition 4.2 gives a constant \( K_1 \) such that for any applicable pair \((A, H)\) there exists a matrix \( \tilde{S}_1 \) satisfying \((A, H) \xrightarrow{\tilde{S}_1} (A_1, \tilde{H}_1)\) for some lower anti–triangular Hankel matrix \( \tilde{H}_1 \), and \( \|I - \tilde{S}_1\| \leq K_1\|A - A_0\|\). Applying Lemma 4.5 with \( T = \tilde{S}_1 \) and \( G = \tilde{H}_1 \), we have that the matrix \( S_1 = a^{-1/2} \tilde{S}_1 \), with \( a = (\tilde{H}_1)_{n,1} \) satisfies the desired bound. Furthermore,

\[
S_1^{-1} AS_1 = (a^{1/2} \tilde{S}_1^{-1})A(a^{-1/2} \tilde{S}_1) = A_1
\]

and

\[
S_1^* HS_1 = (a^{-1/2} \tilde{S}_1^*)H(a^{-1/2} \tilde{S}_1) = a^{-1} \tilde{H}_1 = H_1
\]

or

\[(A, H) \xrightarrow{S_1} (A_1, H_1)\]

by Proposition 4.2, where \( H_1 \) is then of the form (4.7) as desired. \( \square \)

4.2.3. Second step. Zeroing sub–antidiagonal entries of \( H_1 \). Constructing \( S_2 \). With the modified first step, we are now in the same position as before the second step in the model example of this section. The next proposition implements the second step of the model example.

Proposition 4.6 (Zeroing sub–antidiagonal entries of \( H_1 \)). Let \( A_0 \) be a fixed \( H_0 \)-selfadjoint matrix as given in (4.1). Then there exist constants \( K, \delta > 0 \) (depending on \( A_0 \) and \( H_0 \) only) such that for any \( A_1 = J_\mu(\mu) \)

there exists a similitude matrix \( S_2 \) such that

\[(A_1, H_1) \xrightarrow{S_2} (A_1, H_0)\]

satisfying

\[
\|I - S_2\| \leq K (\|A_1 - A_0\| + \|H_1 - H_0\|).
\]

The previous proposition guarantees the existence of a matrix \( S_2 \) that satisfies the following two properties:

1. \( S_2^{-1} A_1 S_2 = A_1 \), and
2. \( S_2^{-1} H_1 S_2 = H_0 \).

That is, it must not affect the matrix \( A_1 \), and it must zero out the sub–antidiagonal elements of the matrix \( H_1 \).
**Lemma 4.7.** Let $R \in \mathbb{R}^{n \times n}$ be a lower anti-triangular Hankel matrix with ones on the main antidiagonal. Then there is a constant $K > 0$ (depending on $R$ only) such that the following statement holds. Denote by $T$ the lower triangular Toeplitz matrix $T = R\tilde{I}$. Write $T = I + E$ and suppose $\|E\| \leq M$ for some bound $M > 1$. Define $S = (f(E))^*$, where

$$f(x) = \sqrt{\frac{1}{1 + x}}.$$

Then

(i) $\|I - S\| \leq K\|E\|$, 
(ii) $S$ is upper triangular Toeplitz,
(iii) $S^*RS = \tilde{I}$.

**Proof.** To prove (i) note that it follows from the definitions that $E$ is nilpotent with index of nilpotence at most $n$. Thus we have $E^n = 0$, and using the power series expansion for $\sqrt{\frac{1}{1 + x}}$ we see that

$$S = f(E) = \sum_{k=0}^{n-1} \left( \prod_{j=1}^{k} \frac{1 - 2j}{2j} \right) E^k$$

and so

$$\|I - S\| \leq \sum_{k=1}^{n-1} \|E\|^k \leq \left( \sum_{k=1}^{n-1} M^{k-1} \right) \|E\| \leq (n - 1) M^{n-2}\|E\|.$$

To prove (ii) note that from (i) $S$ is defined as a transpose of a finite linear combination of powers of a lower triangular Toeplitz matrix $E$, so $S$ is an upper triangular Toeplitz matrix.

(iii) From the definition of $S$, $I = S^*S^*T$, and since $S^*$ and $T$ commute hence

$$I = S^*TS^* = S^*R\tilde{I}S^*.$$

Postmultiplying both side with $\tilde{I}$ gives that

$$\tilde{I} = S^*R\tilde{I}S^*\tilde{I} = S^*RS,$$

completing the proof. \qed

We are now ready to return to the proof of Proposition 4.6.

**Proof of Proposition 4.6.** Letting $R = \epsilon H_1$ where $\epsilon$ is the sign characteristic of the pair $(A_0, H_0)$ and applying Lemma 4.7 (the fact that $\|E\| \leq M$ for some $M$ follows from (4.11), proved in the proof of Lemma 4.5) gives a matrix $S_2$ such that $S_2^*(\epsilon H_1)S_2 = \tilde{I}$, and hence $S_2^*H_1S_2 = H_0$. Since $S_2$ is an upper triangular Toeplitz matrix, it commutes with $A_1$ which is also an upper triangular Toeplitz matrix, and so $S_2^{-1}A_1S_2 = A_1$. These two facts together yield

$$(A_1, H_1) \xrightarrow{S_2} (A_1, H_0).$$

To prove the bound on $\|I - S_2\|$, note that with $R = \epsilon H_1$ and the notations of Lemma 4.7, we have that $\epsilon H_1\tilde{I} = I + E$ which implies $E = \epsilon H_1\tilde{I} - I$. On using the fact that $\|\epsilon\tilde{I}\| = 1$ we have that

$$\|E\| = \|\epsilon\tilde{E}\| = \|H_1 - H_0\| \leq K(\|A - A_0\| + \|H - H_0\|)$$

from (4.11) of the proof of Lemma 4.5. \qed

### 4.2.4. Third step. Combining similitude matrices $S_1$ and $S_2$. We have seen in step 3 of the model example of Section 4.1 that we can combine the similitude matrices $S_1$ and $S_2$ produced by the previous steps to obtain the desired matrix. That is, we already have that

$$(A, H) \xrightarrow{S_1} (A_1, H_1) \xrightarrow{S_2} (A_1, H_0)$$

and hence the matrix $S = S_1S_2$ is such that

$$(A, H) \xrightarrow{S} (A_1, H_0).$$

Moreover, individual Lipschitz-type bounds for $S_1$ and $S_2$ yielded an overall bound of the same form for $S$. In order to accomplish this in general and combine the matrices $S_1$ and $S_2$ of Propositions 4.4 and 4.6, respectively, we will need the following auxiliary result.
Lemma 4.8 (Near-identity similitude matrix yields small perturbations). Let \( A_0 \in \mathbb{C}^{n \times n} \) be an \( H_0 \)-selfadjoint matrix, and \( A \) be an \( H \)-selfadjoint matrix that has the same Jordan structure as \( A_0 \), and \( T \) be an invertible matrix satisfying \((A, H) \xrightarrow{T} (B, G)\) for some \( G \)-selfadjoint matrix \( B \). Suppose that there exist constants \( K, \delta > 0 \) such that
\[
\|I - T\| \leq K (\|A - A_0\| + \|H - H_0\|)
\]
and
\[
\|A - A_0\| + \|H - H_0\| < \delta.
\]
Then
\[
\|G - H_0\| \leq K_4 (\|A - A_0\| + \|H - H_0\|) \quad \text{with} \quad K_4 = (1 + K\delta)^2 + (2 + K\delta)\|H_0\|,
\]
and
\[
\|B - A_0\| \leq K_2 (\|A - A_0\| + \|H - H_0\|) \quad \text{with} \quad K_2 = 2K(1 + K\delta) + 4K\|A_0\|.
\]

Proof. The bound for \( \|G - H_0\| \) is the same as the bound (4.11), and was established in the first part of the proof of Lemma 4.5. To prove the second bound, we have (similar to the proof for \( \|G - H_0\| \)) that
\[
\|B - A_0\| \leq \|T^{-1}\| \cdot \|T\| \cdot \|A - A_0\| + 2\|A_0\| \cdot \|T^{-1}\| \cdot \|I - T\|
\]
by using the obvious identity \( \|I - T^{-1}\| \leq \|T^{-1}\| \cdot \|I - T\| \). Next, by perhaps considering a smaller \( \delta \), we can assume that \( \delta \) is small enough so that \( \|I - T\| \leq \frac{1}{2} \). Then
\[
\|T^{-1}\| - 1 \leq \|I - T^{-1}\| \leq \|T^{-1}\| \cdot \|I - T\| \leq \frac{1}{2} \|T^{-1}\|
\]
which leads to
\[
\|T^{-1}\| \leq 2.
\]
Using this bound for \( \|T^{-1}\| \) in combination with the bound for \( \|T\| \leq M \) of (4.13) in Lemma 4.5, we have
\[
\|B - A_0\| \leq 2KM + 4K\|A_0\| (\|A - A_0\| + \|H - H_0\|).
\]

Theorem 4.9 (Combining similitude matrices \( S_1 \) and \( S_2 \)). Let \( A_0 \in \mathbb{C}^{n \times n} \) be a fixed \( H_0 \)-selfadjoint matrix. There exist positive constants \( K, \delta \) (all depending on \( A_0 \) and \( H_0 \) only) such that for any \( H \)-selfadjoint matrix \( A \) satisfying
\[
\|A - A_0\| + \|H - H_0\| < \delta,
\]
the following statement is true. If we consider mappings
\[
(A, H) \xrightarrow{S_1} (B, G) \xrightarrow{S_2} (C, F)
\]
with any \( S_1, S_2 \) satisfying
\[
\|I - S_1\| \leq K (\|A - A_0\| + \|H - H_0\|), \quad \|I - S_2\| \leq K (\|B - A_0\| + \|G - H_0\|)
\]
then \( S = S_1 S_2 \) satisfies \((A, H) \xrightarrow{S} (C, F)\) and
\[
\|I - S\| \leq (2K + K^2\delta)(\|A - A_0\| + \|H - H_0\|).
\]

Proof. We compute
\[
S^{-1} AS = S_2^{-1} S_1^{-1} AS_1 S_2 = S_2^{-1} BS_2 = C
\]
and
\[
S^* HS = S_2 S_1^* H S_1 S_2 = S_2^* GS_2 = F
\]
using \((A, H) \xrightarrow{S_1} (B, G)\) and then \((B, G) \xrightarrow{S_2} (C, F)\) in each computation. This establishes \((A, H) \xrightarrow{S} (C, F)\). From (4.16) and (4.15),
\[
\|S_i\| \leq 1 + K\delta, \quad i = 1, 2.
\]
Therefore, the matrix \( S = S_1 S_2 \) satisfies
\[
\|I - S\| \leq \|I - S_1\| + \|S_1 S_2\| \leq \|I - S_1\| + (1 + K\delta) \cdot \|I - S_2\| \leq (2K + K^2\delta)(\|A - A_0\| + \|H - H_0\|)
\]
as claimed.
**Proof of Theorem 4.1 in the case of a real single Jordan block.** From Lemma 4.8 we have the individual bounds
\[ \| I - S_i \| \leq K_i (\| A - A_0 \| + \| H - H_0 \|), \quad i = 1, 2 \]
Hence applying Theorem 4.9 yields exactly (4.3), completing the justification of Theorem 4.1 in the case of \((A_0, H_0)\) in the real single Jordan block case. \(\square\)

In the next section, this result is expanded to the single complex eigenvalue case.

### 4.3. Sketch of the proof of Theorem 4.1 in the case of two complex conjugate Jordan blocks

In Section 4.2, Theorem 4.1 was completely proved for the single real Jordan block case. The results need to be modified slightly to complete the proof of Theorem 4.1 and to prove it for the case of a single complex eigenvalue pair\(^4\). Specifically, we now consider the case where
\[
A_0 = \begin{bmatrix} J(\lambda) & 0 \\ 0 & J(\overline{\lambda}) \end{bmatrix}, \quad H_0 = \tilde{I},
\]
where we use the notations of (2.1).

For some pair \((A, H)\) with the same Jordan structure as \((A_0, H_0)\) (that is, \(A\) has Jordan form \(A_1 = J_k(\mu) \oplus J_k(\overline{\mu})\) for some complex \(\mu\)), suppose that a matrix \(S_1\) is found as in Section 4.2.1 such that \((A, H) \xrightarrow{S_1} (A_1, H_1)\).

We demonstrate that the matrix \(H_1\) must have the form
\[
H_1 = \begin{bmatrix} 0 & G^* \\ G & 0 \end{bmatrix}, \quad G = (g_{i+j}) = \begin{bmatrix} 0 & \cdots & 0 & g_{n+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \\ g_{n+1} & g_{n+2} & \cdots & g_{2n} \end{bmatrix},
\]
indeed, from [GLR05, Corollary 4.2.5], it follows that the upper-left and lower-right blocks of \(H_1\) are all zeros as claimed. Further, \(A_1\) is \(H_1\)-selfadjoint, and so \(H_1A_1\) must be selfadjoint and
\[
H_1A_1 = \begin{bmatrix} 0 & GJ(\overline{\mu}) \\ GJ(\mu) & 0 \end{bmatrix},
\]
hence
\[
GJ(\mu) = (G^*J(\overline{\mu}))^* 
\]
and denoting by \(Z\) the lower shift matrix of appropriate size,
\[
GZ^T = ZG 
\]
and so \(G\) is a Hankel matrix as claimed. From this point, Proposition 4.6 can be suitably modified to produce a matrix \(S_2\) such that \((A, H) \xrightarrow{S_1} (A_0, H_1) \xrightarrow{S_2} (A_0, H_0)\) and hence the matrix \(S = S_1S_2\) is as desired.

Theorem 4.1 is now completely proved.

### 5. Proof of Theorem 1.8 for the multiple Jordan block case

In Section 4 the main result, Theorem 1.8, was proved for the case where \((A_0, H_0)\) are in the canonical form of Theorem 2.1 and \(A_0\) was either a single real Jordan block or a direct sum of two conjugate nonreal Jordan blocks. The following theorem generalizes this result to the case where \(A_0\) is an arbitrary Jordan canonical form matrix. In accordance with the Theorem 3.1, this will completely prove the desired Theorem 1.8.

\(^{\text{4Recall that the eigenvalues of } H\text{-selfadjoint matrices are either real or occur in complex conjugate pairs having identical Jordan structure, see, e.g., Theorem 2.1.}}\)
Theorem 5.1 (Extension of Theorem 4.1. Lipschitz stability of similitude matrices in the multiblock case). Let \( A_0 \in \mathbb{C}^{n \times n} \) be a fixed \( H_0 \)-selfadjoint matrix, both \( A_0 \) and \( H_0 \) in the canonical form described in Theorem 2.1. Then there exist constants \( K, \delta > 0 \) (depending on \( A_0 \) and \( H_0 \) only) such that the following assertion holds. For any \( H \)-selfadjoint matrix \( A \) such that \( A \) has the same Jordan structure as \( A_0 \) and

\[
\|A - A_0\| + \|H - H_0\| < \delta,
\]

the pairs \((A_0, H_0)\) and \((A, H)\) are similitude, and there exists a similitude matrix \( S \) such that

\[
\|I - S\| \leq K (\|A - A_0\| + \|H - H_0\|).
\]

This theorem will be proved by induction on the number of the Jordan blocks of \( A_0 \). Theorem 4.1 of Section 4 establishes the result for the case when \( A_0 \) has a single block \( \tilde{J}(\lambda) \). Now, to make the inductive step we will need the following result that allows us to “decouple” such individual blocks from the rest.

Lemma 5.2 (Decoupling). Let \( A_0 \in \mathbb{C}^{n \times n} \) be a fixed \( H_0 \)-selfadjoint matrix, both \( A_0 \) and \( H_0 \) in the canonical form described in Theorem 2.1, and denote

\[
A_0 = \begin{bmatrix} \tilde{J}(\lambda_1) & 0 \\ 0 & \tilde{A}_0 \end{bmatrix} \quad \text{and} \quad H_0 = \begin{bmatrix} \epsilon_1 P_1 & 0 \\ 0 & \tilde{H}_0 \end{bmatrix};
\]

i.e. consider the partition that singles out a block \( \tilde{J}(\lambda_1) \) of the form defined in (2.1). We assume that \( \tilde{J}(\lambda_1) \) is the biggest block of \( A_0 \) corresponding to the eigenvalue \( \lambda_1 \). Then there are positive constants \( K \) and \( \delta \) (depending on \( A_0 \) and \( H_0 \) only) such that for any \( H \)-selfadjoint matrix \( A \) with the same Jordan structure as \( A_0 \) satisfying

\[
\|A - A_0\| + \|H - H_0\| < \delta,
\]

the following statements hold.

(i). \( A \) has an eigenvalue \( \mu_1 \) satisfying

\[
|\lambda_1 - \mu_1| \leq K \|A - A_0\|.
\]

(ii). There exists a similitude matrix \( S \), i.e., \((A, H) \xrightarrow{S} (A_2, H_2)\) such that the matrices \( A_2 \) and \( H_2 \) have the form

\[
A_2 = \begin{bmatrix} \tilde{J}(\mu_1) & 0 \\ 0 & \tilde{A}_1 \end{bmatrix} \quad \text{and} \quad H_2 = \begin{bmatrix} \epsilon_2 P_2 & 0 \\ 0 & \tilde{H}_1 \end{bmatrix},
\]

with some \( \tilde{A}_1, \tilde{H}_1 \) and \( S \) satisfies the bound

\[
\|I - S\| \leq K (\|A - A_0\| + \|H - H_0\|).
\]

The proof of this lemma is postponed to Section 7. Lemma 5.2 allows us to make an inductive step and to prove Theorem 5.1.

Proof of Theorem 5.1. Applying Lemma 5.2 to the pair \((A, H)\), as in

\[
(A, H) \xrightarrow{S_1} (A_1, H_1),
\]

results in a pair \((A_1, H_1)\) of the form shown in (5.4) with

\[
\|I - S_1\| \leq K_1 (\|A - A_0\| + \|H - H_0\|).
\]

Since the lower right block \( \tilde{A}_1 \) of \( A_1 \) has less Jordan blocks than \( A_0 \), we can use the inductive assumption and apply Theorem 5.1 to the pair \((A_1, H_1)\), as in

\[
(A_1, H_1) \xrightarrow{S_2} (A_2, H_0),
\]

with

\[
\|I - S_2\| \leq K_2 (\|A_1 - A_0\| + \|H_1 - H_0\|).
\]

From this follows

\[
\|I - S_2\| \leq K_2 (\|A_2 - A_0\| + \|H_2 - H_0\|)
\]

with

\[
S_2 = \begin{bmatrix} I & 0 \\ 0 & S_2 \end{bmatrix}.
\]
Indeed, the transition from (5.7) to (5.8) involves only adding the $(1, 1)$ blocks into consideration, and from (5.2) and (5.4), we see that the only nontrivial addition is from $\|J(\mu_1) - J(\lambda_1)\|$. In view of Lemma 6.14 (which is proved independently of results of this theorem), this serves only to modify the constant, and (5.8) is established. Observe, that from (6.33) and (5.7) it follows that $K'_2 = \max\{1, K_2\}$ depends on $A_0$ and $H_0$ only.

Finally, Theorem 4.9 allows us to combine the bounds (5.6) and (5.8), and it implies the desired bound (5.5) for $S = S_1 S_2$. $\square$

In order to complete the proof of Theorem 5.1 (and thus, of the main result, Theorem 1.8), it remains only to prove Lemma 5.2. It will be done in Section 7 after we obtain in the next auxiliary section several necessary results on the perturbation of subspaces.

6. Auxiliary lemmas on semigaps, gaps and perturbations of subspaces

In order to provide the proof of Lemma 5.2 in Section 7 one needs to use a number of results on small perturbations of certain invariant subspaces. We have gathered all these auxiliary results in this section. In order to obtain specific bounds for such perturbations, we need to deal with the distance between two subspaces. One standard way to define such a distance is based on the concept of a gap. Our approach below is slightly different, it is based on the concept of a semigap for which it is often easier to obtain the desired bounds. The key result of this section is Lemma 6.6 that says that when two subspaces have equal dimensions the gap between them is equal to the semigap. Many of the auxiliary results in this section are known. For somewhat less know results, e.g., about semigaps, we provide references and in some cases new proofs.

6.1. Gap between subspaces

Before defining the concept of a gap, let us recall that

• A matrix $P_M$ is called a projector onto a subspace $M \subset \mathbb{C}^n$ if (a) $\text{Im} P_M = M$; (b) $P_M^2 = P_M$.
• Further, $P_M$ is called an orthogonal projector onto $M$ if additionally we have (c) $P_M^* = P_M$.

Here is the key definition.

Definition 6.1 (Gap. First definition). The gap $\theta(M, N)$ between two subspaces $M, N \subset \mathbb{C}^n$ can be introduced via

$$\theta(M, N) = \|P_M - P_N\|$$

where $P_M$ denotes the orthogonal projector onto $M$.

It is well-known, see, e.g., [GLR86], that gap is a metric in the set of all subspaces of $\mathbb{C}^n$.

The definition in (6.1) has been found to useful in many instances. However, in Theorem 2.6 and Lemma 4.3 above we did not deal with entire subspaces, but rather with particular vectors spanning them. Therefore, in our context, it is often more convenient to use the next definition that is well-known to be equivalent to the first one.

Definition 6.2 (Gap. Second definition). The gap $\theta(M, N)$ between two subspaces $M, N \subset \mathbb{C}^n$ can be introduced via

$$\theta(M, N) = \max \left\{ \sup_{x \in M} \inf_{y \in N} \|x - y\|, \sup_{y \in N} \inf_{x \in M} \|y - x\| \right\}$$

Theorem 2.6 and Lemma 4.3 have just been used as a motivation for (6.2), but in these statements the vectors $\{f_k\}$ were fixed, while the vectors $\{g_k\}$ were their perturbations. One might expect that in some instances it might be easier to rely on the properties of fixed vectors $\{f_k\}$, but the quantity $\theta(\text{span}\{f_k\}, \text{span}\{g_k\})$ is clearly symmetric, so it does not give $\{f_k\}$'s any “advantage.” In order to better capture the difference between $\{f_k\}$’s and $\{g_k\}$’s we now introduce a different one-sided quantity.
Lemma 6.6 (Semigaps for subspaces of the same dimension).

The dimensions of $\dim M$ and $\dim N$ are called the semigap (or one-sided gap) from $M$ to $N$.

Proof. In order to prove (6.4) we need to consider three cases.

Case 1. $\dim M = \dim N = 1$: If $M = N$ there is nothing to prove, so we assume $M \neq N$.

Let us choose $x \in N$ and $y \in M$ such that $\|x\| = \|y\| = 1$. By an appropriate unimodular rescaling of $y$ one can guarantee that $x^* y \in \mathbb{R}$. Further, let us define

$$U = I - 2 wy^*, \quad \text{where} \quad w = \frac{1}{\|x - y\|} (x - y).$$

It is well-known (and can be easily checked) that the Householder reflection $U$ is Hermitian and unitary. In particular,

$$U^2 = I.$$  \hfill (6.5)

Secondly, it can be easily checked that

$$Ux = y, \quad Uy = x.$$  \hfill (6.6)

Example 6.5 (Semigaps can be nonsymmetric). Let $N = \mathbb{C}^2$ and $M = \text{span}\{e_1\}$. Since in this case $M \subset N$ hence

$$\theta_0(M, N) = 0.$$

However, the vector $e_2 \in N$ is orthogonal to $M$ and hence

$$\theta_0(N, M) = 1.$$

In the above example the dimensions of $M$ and $N$ were different. The next statement shows that when the dimensions of $M$ and $N$ are the same then the two associated semigaps are always equal.

Lemma 6.6 (Semigaps for subspaces of the same dimension). Let $M, N \subset \mathbb{C}^n$ be two subspaces. If $\dim M = \dim N$, then

$$\theta_0(M, N) = \theta_0(N, M) \quad (= \theta(M, N)).$$  \hfill (6.4)

Proof. In order to prove (6.4) we need to consider three cases.

Case 1. $\dim M = \dim N = 1$: If $M = N$ there is nothing to prove, so we assume $M \neq N$.

• Defining an appropriate (complex) Householder reflection. Let us choose $x \in N$ and $y \in M$ such that $\|x\| = \|y\| = 1$. By an appropriate unimodular rescaling of $y$ one can guarantee that $x^* y \in \mathbb{R}$. Further, let us define

$$U = I - 2 wy^*, \quad \text{where} \quad w = \frac{1}{\|x - y\|} (x - y).$$

It is well-known (and can be easily checked) that the Householder reflection $U$ is Hermitian and unitary. In particular,

$$U^2 = I.$$  \hfill (6.5)

Secondly, it can be easily checked that

$$Ux = y, \quad Uy = x.$$  \hfill (6.6)
Indeed,
\[ Ux = x - 2\frac{(y - x)(y - x)^*}{(y - x)^*(y - x)} x = x - \frac{(x - y)(x^*y - 1)}{y^*x - 1} = y. \]

The second equation in (6.6) follows from the first one and (6.5).

Finally, observe
\[ U P_N U = P_M. \]  \hfill (6.7)

- **Proving** $\theta_0(\mathcal{M}, \mathcal{N}) = \theta_0(\mathcal{N}, \mathcal{M})$. From (6.6) and (6.7) and from the fact that $U$ is unitary it follows that
\[ \|x - P_N x\| = \|U y - U P_N U y\| = \|U(y - P_M y)\| = \|y - P_M y\|. \]

This and the property (6.3a) imply the desired (6.4).

**Case 2.** $\dim \mathcal{M} = \dim \mathcal{N} = k > 1$ and $\theta_0(\mathcal{M}, \mathcal{N}) = 1$: First we observe that
\[ \theta_0(\mathcal{M}, \mathcal{N}) = 1 \iff \mathcal{M} \cap \mathcal{N} = \{0\} \]  \hfill (6.8)

and
\[ \theta_0(\mathcal{N}, \mathcal{M}) = 1 \iff \mathcal{N} \cap \mathcal{M} = \{0\} \]  \hfill (6.9)

and our task is to show that (6.8) implies (6.9).

Denote by $\mathcal{M}_1$ the orthogonal complement to $\mathcal{M} \cap \mathcal{N}$ in $\mathcal{M}$ and define $\mathcal{N}_1 = P_N \mathcal{M}_1$. In view of (6.8) we have $\dim \mathcal{M}_1 < \dim \mathcal{M}$ and hence $\dim \mathcal{N}_1 < \dim \mathcal{N}$. Therefore there exist $y \in \mathcal{N}$ that is orthogonal to $\mathcal{N}_1$. Clearly, this $y$ is orthogonal to $\mathcal{M}$ implying (6.9).

**Case 3.** $\dim \mathcal{M} = \dim \mathcal{N} = k > 1$ and $\theta_0(\mathcal{M}, \mathcal{N}) < 1$: In order to prove (6.4) we first observe that
\[ P_M N = \mathcal{M}, \quad P_N M = \mathcal{N}. \]  \hfill (6.10)

Indeed, the result proved in case 2 above implies that if $\theta_0(\mathcal{M}, \mathcal{N}) < 1$ we must also have $\theta_0(\mathcal{N}, \mathcal{M}) < 1$. Therefore none of $\mathcal{M}, \mathcal{N}$ contains vectors orthogonal to each other and (6.10) follows.

Now, since the unit circle in $\mathcal{M}$ is compact hence the supremum in (6.3a) is attained, i.e, there exists $x \in \mathcal{M}$ such that $\|x\| = 1$, we have that
\[ \theta_0(\mathcal{M}, \mathcal{N}) = \|x - P_N x\|. \]

Denoting $\mathcal{M}_x = \text{span}\{x\}$ we have
\[ \theta_0(\mathcal{M}_x, \mathcal{N}) = \theta_0(\mathcal{M}, \mathcal{N}). \]  \hfill (6.11)

Now, in view of (6.10) there is a subspace $\mathcal{N}_x \subset \mathcal{N}$ such that
\[ \mathcal{M}_x = P_M \mathcal{N}_x. \]

Due to this particular choice of $\mathcal{N}_x$ we have
\[ \theta_0(\mathcal{N}_x, \mathcal{M}_x) = \theta_0(\mathcal{N}_x, \mathcal{M}). \]  \hfill (6.12)

Since both $\mathcal{M}_x$ and $\mathcal{N}_x$ are one-dimensional, we have
\[ \theta_0(\mathcal{M}_x, \mathcal{N}_x) = \theta_0(\mathcal{N}_x, \mathcal{M}_x) \]  \hfill (6.13)

Gathering all the above we have
\[ \theta_0(\mathcal{M}, \mathcal{N}) \overset{(6.11)}{=} \theta_0(\mathcal{M}_x, \mathcal{N}) \overset{(6.3b)}{\leq} \theta_0(\mathcal{M}_x, \mathcal{N}_x) \overset{(6.12)}{=} \theta_0(\mathcal{N}_x, \mathcal{M}_x) \overset{(6.3b)}{=} \theta_0(\mathcal{N}_x, \mathcal{M}) \leq \theta_0(\mathcal{N}, \mathcal{M}) \]

We have proved that $\theta_0(\mathcal{M}, \mathcal{N}) \leq \theta_0(\mathcal{N}, \mathcal{M})$ without making any assumptions on $\mathcal{M}$ and $\mathcal{N}$. Hence the desired result (6.4) follows by symmetry.

The proof is complete. \hfill \Box

**Corollary 6.7.** If $\dim \mathcal{M} = \dim \mathcal{N}$ then
\[ \theta(\mathcal{M}, \mathcal{N}) = \sup_{x \in \mathcal{M}} \|x - P_N x\| = \sup_{x \in \mathcal{M}} \inf_{y \in \mathcal{N}} \|x - y\|. \]

In [O91] the above lemma and corollary were found to be useful to study the change of Jordan structure of $H$-selfadjoint matrices under small perturbations.
Lemma 6.8. Let $\mathcal{M} = \text{span}\{f_i\}_{i=0}^{m-1}$ where vectors $\{f_i\}$ are linearly independent. There exists a constant $K > 0$ (depending on $\{f_i\}$ only) such that for any set of vectors $\{g_i\}_{i=0}^{m-1}$ satisfying

$$\|f_i - g_i\| \leq K,$$

we have

$$\theta_0(\mathcal{M}, \mathcal{N}) \leq K\|P_{\mathcal{F} - \mathcal{E}}\|,$$

(6.14)

where $\mathcal{N} = \text{span}\{g_i\}$, and $P_{\mathcal{F} - \mathcal{E}}$ denotes the change-of-coordinates matrix from the standard basis $E = \{e_i\}$ to $F = \{f_i\}$.

Proof. Let $x \in \mathcal{M}$ and $\|x\| = 1$. For the decomposition of $x$ with respect to the fixed basis $F\{f_i\}$:

$$x = \alpha_0f_0 + \ldots + \alpha_{m-1}f_{m-1},$$

let us consider

$$y = \alpha_0g_0 + \ldots + \alpha_{m-1}g_{m-1}.$$

Clearly, $y \in \mathcal{N}$ and

$$\|x - y\| \leq \alpha_0\|f_0 - g_0\| + \ldots + \alpha_{m-1}\|f_{m-1} - g_{m-1}\| \leq \max_{0 \leq k \leq m-1} |\alpha_k| \cdot mK. \quad (6.15)$$

In order to complete the proof we need to find a bound on $\max_{0 \leq k \leq m-1} |\alpha_k|$. To this end let us consider two decompositions of $x$ with respect to the fixed basis $F = \{f_i\}$ and the standard (also fixed) basis $E = \{e_i\}$, respectively:

$$x = \alpha_1f_1 + \ldots + \alpha_nf_n = x_1e_1 + \ldots + x_ne_n,$$

i.e.,

$$x = [x]_E = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad [x]_F = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Now,

$$\max_{0 \leq k \leq m-1} |\alpha_k| \leq \sqrt{|\alpha_0|^2 + \ldots + |\alpha_{m-1}|^2} = \|x\|_F \leq \|P_{\mathcal{F} - \mathcal{E}} \cdot [x]_E\| \leq \|P_{\mathcal{F} - \mathcal{E}}\|$$

(6.16)

and we see that (6.16) and (6.15) imply (6.14). \qed

6.3. Bounds for the perturbations of $H$-orthogonal companions

The following definition [GLR83] introduces a counterpart $\mathcal{M}^{\perp n}$ (or simply $\mathcal{M}^{\perp}$) of the usual orthogonal complement $\mathcal{M}^{\perp}$ for the spaces with indefinite inner product induced by $H$.

Definition 6.9 ($H$-orthogonal companion). Let $H \in \mathbb{C}^n$ be an invertible Hermitian matrix. For a subspace $\mathcal{M} \subset \mathbb{C}^n$ its $H$-orthogonal companion is defined as

$$\mathcal{M}^{\perp} = \{x \in \mathbb{C}^n : \langle x, y \rangle_H = 0 \ \forall y \in \mathcal{M}\}.$$

In Section 7 we will need to develop an inductive decoupling process passing from a given subspace $\mathcal{M}$ and its small perturbation $\mathcal{N}$ to their orthogonal companions $\mathcal{M}^{\perp}$ and $\mathcal{N}^{\perp}$, respectively. The following lemma will be key in this framework since it claims that in this case the subspace $\mathcal{M}^{\perp}$ is a small perturbation of $\mathcal{N}^{\perp}$.

Lemma 6.10 (Bounds for $H$-orthogonal companions). Let $H_0$ be a fixed invertible Hermitian matrix, and let $\mathcal{M} \subset \mathbb{C}^n$ be a fixed subspace. There is a constant $\delta > 0$ (depending on $H_0$ and $\mathcal{M}$ only) such that for any invertible Hermitian $H$ satisfying

$$\|H - H_0\| \leq \delta,$$

and any subspace $\mathcal{N}$ satisfying

$$\dim \mathcal{M} = \dim \mathcal{N}, \quad \theta_0(\mathcal{M}, \mathcal{N}) \leq L \quad (6.17)$$

(with certain $L$) we have

$$\theta(\mathcal{M}^{\perp}, \mathcal{N}^{\perp}) \leq (\kappa(H_0) + 2\|H_0\|) \cdot (L + \|H - H_0\|). \quad (6.18)$$
Proof. First observe that $H^{-1}P_{M^+}H$ is a projection onto subspace $H^{-1}M^+$, though not necessarily orthogonal. Second, it is well known [GLR05] that for two (possibly not orthogonal) projectors $Q_M$ and $Q_N$ onto $M$ and $N$, respectively, we have
\[
\|P_M - P_N\| \leq \|Q_M - Q_N\|.
\]
Third, it is well-known [GLR05] that
\[
M^{[1]} = H_0^{-1}M^+, \quad N^{[1]} = H^{-1}N^+.
\]
From the above three facts it follows that
\[
\theta(M^{[1]}, N^{[1]}) = \|P_{H_0^{-1}M^+} - P_{H^{-1}N^+}\| \leq \|H_0^{-1}P_{M^+}H_0 - H^{-1}P_{N^+}H\|. \tag{6.19}
\]
In order to bound the latter let us proceed with the right hand side of (6.19), and add and subtract the quantity $H_0^{-1}P_{N^+}H_0$ and $H^{-1}P_{N^+}H$:
\[
\|H_0^{-1}P_{M^+}H_0 - H^{-1}P_{N^+}H_0\| + (H_0^{-1}P_{N^+}H_0 - H_0^{-1}P_{N^+}H) + (H_0^{-1}P_{N^+}H - H^{-1}P_{N^+}H)\]
\[
\leq \kappa(H_0)\|P_{M^+} - P_{N^+}\| + \|H_0^{-1}\|\|H - H_0\| + \|H_0^{-1}\|\|H - H_0\|, \tag{6.20}
\]
where $\kappa(H_0) = \|H_0\|\|H_0^{-1}\|$. Combining (6.19) and (6.20) we finally obtain
\[
\theta(M^{[1]}, N^{[1]}) \leq \kappa(H_0) \cdot \theta(M^+, N^+) + 2\|H_0^{-1}\|\|H - H_0\| \tag{6.21}
\]
Since
\[
\|P_{M^+} - P_{N^+}\| = \|(I - P_M) - (I - P_N)\| = \|P_M - P_N\|
\]
hence
\[
\theta(M^+, N^+) = \theta(M, N),
\]
which together with (6.21) and (6.17) imply the desired (6.18). \qed

6.4. Several useful bounds

Let $\lambda$ be an eigenvalue of $A_0 \in \mathbb{C}^{n \times n}$. Recall that the root subspace $\mathcal{R}(A_0, \lambda)$ is defined as a linear span of all Jordan chains of $A_0$ corresponding to $\lambda$, see, e.g., [GLR66]. Alternatively, $\mathcal{R}(A_0, \lambda) = \text{Ker}(A_0 - \lambda I)^n$. Clearly, the dimension of $\mathcal{R}(A_0, \lambda)$ is equal to the total algebraic multiplicity of the eigenvalue $\lambda$ of $A_0$.

Further, let $\Gamma$ be a simple (without self-intersections), closed, rectifiable contour with no eigenvalues of $A_0$ on it. Let $\{\lambda_1, \ldots, \lambda_r\}$ be a set of all eigenvalues of $A_0$ inside $\Gamma$. Denote
\[
\mathcal{R}(A_0, \Gamma) = \mathcal{R}(A_0, \lambda_1) + \ldots + \mathcal{R}(A_0, \lambda_r).
\]
With these notations the following bound holds.

Lemma 6.11 (Bound for the perturbation of root subspaces). Let $A_0 \in \mathbb{C}^{n \times n}$, and let $\Gamma$ be a simple, closed, rectifiable contour such that $A_0$ does not have eigenvalues on $\Gamma$. Then there are constants $K_{\text{root}}, \delta > 0$ (depending on $A_0$ and $\Gamma$ only) such that any matrix $A$ satisfying $\|A - A_0\| \leq \delta$ and does not have any eigenvalues on $\Gamma$, then
\[
\theta(\mathcal{R}(A_0, \Gamma), \mathcal{R}(A, \Gamma)) \leq K_{\text{root}}\|A - A_0\|. \tag{6.22}
\]
In particular, the total multiplicity of all eigenvalues inside $\Gamma$ is the same for $A_0$ and $A$.

The proof of the latter lemma can be found, e.g., in [GLR05], p. 334. We will also need the following result (cf., e.g., with [O91]).

Lemma 6.12 (Bound for adjusting matrix $S$). Let the decomposition
\[
\mathbb{C}^n = M_1 + \ldots + M_k
\]
be given. For any decomposition
\[
\mathbb{C}^n = N_1 + \ldots + N_k
\]
with
\[
\dim M_j = \dim N_j, \quad j = 1, 2, \ldots, k,
\]
there exists an invertible matrix $S \in \mathbb{C}^{n \times n}$ (that we suggest to call the adjusting matrix) satisfying
\[
S M_j = N_j \tag{6.26a}
\]
\[
\|I - S\| \leq \sum_{j=1}^k \theta(M_j, N_j). \tag{6.26b}
\]
Proof. The proof is presented by considering two cases.
Case 1. $\theta(\mathcal{M}_j, \mathcal{N}_j) < 1$, for $j = 1, 2, \ldots, k$: It is easy to verify that the matrix $S$ defined as

$$S = I - \sum_{j=1}^{k} (P_{\mathcal{M}_j} - P_{\mathcal{N}_j}),$$  \hfill (6.27)

satisfy

$$SM_j \subset \mathcal{N}_j.$$  \hfill (6.28)

Secondly, $S$ is invertible. Indeed, for any nonzero $x \in \mathbb{C}^n$, let $x = x_1 + x_2 + \cdots + x_k$ where $x_j \in \mathcal{M}_j$ we have

$$Sx = P_{\mathcal{N}_j}x_1 + \cdots + P_{\mathcal{N}_j}x_k.$$  

Since the $\theta(\mathcal{M}_j, \mathcal{N}_j) < 1$, hence $\mathcal{M}_j$ and $\mathcal{N}_j$ are not orthogonal. Therefore, if $x_j \neq 0$ then $P_{\mathcal{N}_j}x_j \neq 0$. This implies $\ker S = \{0\}$ so that $S$ is invertible. In view of (6.25), the invertibility of $S$ and (6.28) imply (6.26a). For $S$ defined by (6.27) the relation (6.26b) is obvious:

$$\|I - S\| \leq \sum_{j=1}^{k} \|P_{\mathcal{N}_j} - P_{\mathcal{M}_j}\| = \sum_{j=1}^{k} \theta(\mathcal{M}_j, \mathcal{N}_j).$$

Case 2. $\theta(\mathcal{M}_j, \mathcal{N}_j) = 1$, for some $j$: In view of (6.23), (6.24) and (6.25) we can always choose an invertible matrix $T$ such that $T\mathcal{M}_j = \mathcal{N}_j$. Setting $S = (||T||)^{-1}T$ we have

$$\|I - S\| \leq \|I\| + \|S\| \leq 2.$$  

This and the fact that $1 \leq \sum_{j=1}^{k} \theta(\mathcal{M}_j, \mathcal{N}_j)$ imply the desired (6.26b). \hfill \Box

In order to obtain necessary bounds on the perturbation of the eigenvalues of matrices we will need the following auxiliary result.

Lemma 6.13. Let $A_0 \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{C}$, and $a_1, a_2 \in \mathbb{C}^n$ satisfy $(A_0 - \lambda I)a_2 = a_1$. There is a constant $\delta > 0$ (depending on $A_0$, $a_1$, and $a_2$ only) such that for any $A \in \mathbb{C}^{n \times n}$, $\mu \in \mathbb{C}$, and $b_1, b_2 \in \mathbb{C}^n$ satisfying $(A - \mu I)b_2 = b_1$ and

$$\|A - A_0\| \leq \delta, \quad |\mu - \lambda| \leq K_{\text{eig}}\|A - A_0\|, \quad \|b_2 - a_2\| \leq K_{\text{vec}}\|A - A_0\|$$  \hfill (6.29)

with some $K_{\text{eig}}, K_{\text{vec}} > 0$ we have

$$\|b_1 - a_1\| \leq K_{\text{next}}\|A - A_0\|$$  \hfill (6.30)

with

$$K_{\text{next}} = \|A_0\| + \delta + \|a_2\|(K_{\text{eig}} + 1) + (K_{\text{eig}}\delta + |\lambda|)K_{\text{vec}}$$  \hfill (6.31)

Proof. It follows from the first two inequalities in (6.29) that

$$\|A\| \leq \|A_0\| + \delta, \quad |\mu| \leq K_{\text{eig}}\delta + |\lambda|.$$  

Using this and all the bounds in (6.29) we have

$$\|b_1 - a_1\| = \|(A - \mu I)b_2 - (A_0 - \lambda I)a_2\| = \|Ab_2 - A_0a_2 + Aa_2 - A_0a_2 + \lambda a_2 - \mu a_2 + \mu a_2 - \mu b_2\|$$

$$\leq \|A\| \cdot \|b_2 - a_2\| + \|a_2\| \cdot \|A - A_0\| + \|a_2\| \cdot |\mu - \lambda| + |\mu| \cdot \|b_2 - a_2\|$$

$$\leq ((\|A_0\| + \delta + \|a_2\|(K_{\text{eig}} + 1) + (K_{\text{eig}}\delta + |\lambda|)K_{\text{vec}})\|A - A_0\|,$$

and (6.30) follows. \hfill \Box

Finally, here is the second key lemma of this section to be used in the proof of Lemma 5.2 in Section 7.

Lemma 6.14 (Perturbations of the eigenvalues and of Jordan chains). Let $\Gamma$ be a simple, closed, rectifiable contour such that $A_0$ does not have eigenvalues on $\Gamma$. Let $\sigma(A_0) = \{\lambda_1, \ldots, \lambda_\gamma\}$ be the set of all eigenvalues of $A_0 \in \mathbb{C}^{n \times n}$ inside $\Gamma$. There are constants $K, \delta > 0$ (depending on $A_0$ only) such that the following statements holds. For any $A \in \mathbb{C}^{n \times n}$ satisfying

$$\|A - A_0\| \leq \delta$$  \hfill (6.32)

we have the following.

(i). If $A$ has exactly $\gamma$ eigenvalues inside $\Gamma$, there is a certain ordering $\{\mu_1, \ldots, \mu_\gamma\}$ of them such that

$$|\lambda_i - \mu_i| \leq K\|A - A_0\|, \quad i = 1, 2, \ldots, \gamma.$$  \hfill (6.33)
(ii). Let \( \{f_k\}_{k=0}^{m-1} \) be the longest Jordan chain corresponding to the eigenvalue \( \lambda_i \in \sigma_i(A_0) \). If the maximal length of the Jordan chain corresponding to its eigenvalue \( \mu_k \) described in (6.33) is also \( m \), then there exists a Jordan chain \( \{g_k\}_{k=0}^{m-1} \) of \( A \) corresponding to \( \mu_i \) such that
\[
\|f_k - g_k\| \leq K\|A - A_0\|, \quad k = 0, \ldots, m-1.
\]

**Proof.** Here is the proof of the two parts of the lemma.

(i). For \( i = 1, 2, \ldots, \gamma \) let \( \Gamma_i \) denote a small circle that contains only one eigenvalue \( \lambda_i \) of \( A_0 \). In accordance with Lemma 6.11 there is \( \delta > 0 \) such that any matrix \( A \) satisfying \( \|A - A_0\| \leq \delta \) will have at least one eigenvalue inside \( \Gamma_i \) (indeed, the total multiplicity of the eigenvalues inside each \( \Gamma_i \) is preserved). Since \( A \) has exactly \( \gamma \) eigenvalues hence it must have exactly one eigenvalue, say, \( \mu_i \) inside each \( \Gamma_i \). Using Lemma 6.11 again we see that there are constants \( K_i > 0 \) (depending only on \( A_0 \)) such that
\[
\theta(\mathcal{R}(A_0, \lambda_i), \mathcal{R}(A, \mu_i)) \leq K_i\|A - A_0\|, \quad i = 1, \ldots, \gamma.
\]

Denote
\[
\mathcal{M}_i = \mathcal{R}(A_0, \lambda_i), \quad \mathcal{N}_i = \mathcal{R}(A, \mu_i), \quad i = 1, \ldots, \gamma,
\]
and
\[
\mathcal{M}_{\gamma+1} = \mathcal{N}_{\gamma+1} = (\mathcal{M}_1 + \cdots + \mathcal{M}_\gamma)^{\perp}.
\]

By Lemma 6.12 there is an \( S \) satisfying
\[
S\mathcal{R}(A_0, \lambda_i) = \mathcal{R}(A, \mu_i), \quad i = 1, \ldots, \gamma,
\]
and
\[
\|I - S\| \leq \sum_{i=1}^\gamma \theta(\mathcal{R}(A_0, \lambda_i), \mathcal{R}(A, \mu_i)).
\]

Combining the latter two bounds (6.35) and (6.37) one obtains
\[
\|I - S\| \leq K_0\|A - A_0\|
\]
with \( K_0 = K_1 + \cdots + K_\gamma \).

Now, let matrix \( R \) be such that
\[
R^{-1}A_0R = \begin{bmatrix}
A^{(0)}_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & A^{(0)}_\gamma & * \\
0 & \cdots & 0 & A^{(0)}_{\gamma+1}
\end{bmatrix}
\]
where \( A^{(0)}_1 \) has the only eigenvalue \( \lambda_i \in \sigma_1(A_0) \), and all the eigenvalues of \( A^{(0)}_{\gamma+1} \) are outside of \( \Gamma \).

Observe that \( R \) depends on \( A_0 \) only, and it can be fixed in advance. The property (6.36) yields that \( R \) diagonalizes \( A_1 = S^{-1}AS \) as well, and moreover,
\[
R^{-1}A_1R = \begin{bmatrix}
A^{(1)}_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & A^{(1)}_\gamma & * \\
0 & \cdots & 0 & A^{(1)}_{\gamma+1}
\end{bmatrix}
\]
with each \( A^{(1)}_j \) having only one eigenvalue \( \mu_i \), and with \( A^{(1)}_{\gamma+1} \) having only the eigenvalues outside of \( \Gamma \).

Further, denoting the size of \( A^{(0)}_1 \) by \( r_i \) we have
\[
|\lambda_i - \mu_i| \leq \frac{1}{r_i} |\text{trace}(A^{(0)}_1) - \text{trace}(A^{(1)}_1)| \leq \frac{1}{r_i} \sum_{k=1}^{r_i} e_k^T (A^{(0)}_1 - A^{(1)}_1)e_k \leq \|A^{(0)}_1 - A^{(1)}_1\| \leq \|R^{-1}A_0R - R^{-1}A_1R\| \leq \kappa(R)\|A_0 - A_1\| = \kappa(R)\|A_0 - S^{-1}AS\| \leq K\|A_0 - A\|.
\]

In the above chain of inequalities the latter one is deduced from (6.38) using the arguments identical to those of the proof of Lemma 4.8. This concludes the proof of the part (i).
(ii). Let us define
\[ g_{m-1} = P_{\mathbb{R}(A)} f_{m-1}, \quad g_{k-1} = (A - \mu_i I) g_k \quad k = 1, 1, \ldots, m - 1. \]

It follows from (6.35) that
\[ \| f_{m-1} - g_{m-1} \| = \| f_{m-1} - P_{\mathbb{R}(A)} f_{m-1} \| \leq \theta(A_0, \lambda), \quad \| f_{m-1} \| \leq K_{vec} \| A - A_0 \|, \]

where \( K_{vec} = \| f_{m-1} \| K_t. \) Lemma 6.13 implies that
\[ \| f_{m-2} - g_{m-2} \| \leq K_{next} \| A - A_0 \|, \]

with \( K_{next} \) given by (6.31). It is easy to see from (6.31) that since \( K_{vec} \) depends on \( A_0 \) and the choice of the (fixed) chain \( \{f_k\} \) hence the constant \( K_{next} \), while possibly bigger, has the same property. Applying the same arguments recursively to \( f_{m-2}, g_{m-2} \), then to \( f_{m-3}, g_{m-3} \), and so on, one obtains, after \( m \) steps, the desired bound (6.34), in which the constant \( K \) is the maximum of the constants obtained in each of these steps.

In order to complete the proof of (ii) we need to show that \( \{g_k\}_{k=0}^{m-1} \) is indeed a Jordan chain. To this end we need to show two things.

First, the vectors \( \{g_k\}_{k=0}^{m-1} \) have to be linearly independent. Since \( \delta > 0 \) is at our disposal we may assume it to be small enough so that the bound (6.32), linear independence of \( \{f_k\} \) and (6.34) guarantee linear independence of \( \{f_k\} \).

Secondly, we need to show that \( (A - \mu I) g_0 = 0 \). This follows form our assumption that the longest Jordan chain of \( A \) corresponding to \( \mu \) has length \( m \).

\[ \square \]

7. Proof of Lemma 5.2

In this section we prove Lemma 5.2 which completes the proof of Theorem 5.1 and thus of the main result, Theorem 1.8. As before, we start with a clarifying example.

7.1. Multiple Block Model example

For some \( \delta > 0 \), consider the matrices \( A_0 \) and \( H_0 \) defined by
\[ A_0 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad H_0 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \]

and small perturbations of the above matrices \( A \) and \( H \) as,
\[ A = \begin{bmatrix} 0 & 1 & 2\delta & 0 & \delta \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \delta \\ 1 & 0 & 2\delta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & \delta & 0 & 1 & 0 \end{bmatrix}, \]

We seek a matrix \( S \) such that
\[ (A, H) \xrightarrow{S} (A_1, H_1), \quad \|I - S\| \leq K(\|A - A_0\| + \|H - H_0\|) \]

where \( A_1 \) and \( H_1 \) have the forms
\[ A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & * & * \end{bmatrix} \quad \text{and} \quad H_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * \end{bmatrix} \].
that is, the matrix $S$ decouples the first Jordan chain from the remaining ones. Such a process in general enables us to proceed inductively on the later portions. As in Section 4.1 the matrix $S$ is found in two steps, the first is such that
\[ (A, H) \xrightarrow{S_1} (A_1, H'_1) \quad \text{and} \quad \|I - S_1\| \leq K\|A - A_0\| + \|H - H_0\| \]
where $H'_1$ has the form
\[
H'_1 = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 1 & * & 0 & 0 \\
1 & * & * & 0 & 0 \\
0 & 0 & 0 & * & * \\
0 & 0 & 0 & * & *
\end{bmatrix}.
\]
Then, in the second step, we produce a matrix $S_2$ such that
\[ (A_1, H'_1) \xrightarrow{S_2} (A_1, H_1) \quad \text{and} \quad \|I - S_2\| \leq K\|A - A_0\| + \|H - H_0\|). \]
Finally, we show that $S = S_1S_2$ satisfies
\[ (A, H) \xrightarrow{S} (A_1, H_1) \quad \text{and} \quad \|I - S\| \leq K\|A - A_0\| + \|H - H_0\|) \]
as desired.

### 7.1.1 First step. Mapping $A \rightarrow A_1$, Constructing $S_1$

As a first attempt, we will try to proceed as in the example in Section 4.1 and choose a similarity matrix $S_1$ that maps the first Jordan chain of $A_0$,

\[ 0 \leftarrow e_1 \leftarrow e_2 \leftarrow e_3, \]

to that of $A$,

\[ 0 \leftarrow e_1 \leftarrow e_2 + (2\delta)e_1 \leftarrow e_3. \]

This attempt will fail, but it will indicate a difficulty and a way to resolve it. Denote these vectors by $g_k$, 

\[
g_3 = e_3, \quad g_2 = e_2 + (2\delta)e_1, \quad g_1 = e_1.
\]

As before we choose an $S_1$ such that $S_1: e_k \rightarrow g_k$ for $k = 1, 2, 3$, but this leaves a choice of where to map the vectors in root subspaces corresponding to later Jordan chains, in this case $e_4$ and $e_5$. As a first attempt, let us choose $S_1$ to leave these unchanged. So we initially choose to map them as follows:

\[
\begin{array}{c}
S_1: e_1 \rightarrow g_1 \\
S_1: e_2 \rightarrow g_2 \\
S_1: e_3 \rightarrow g_3 \\
S_1: e_4 \rightarrow e_4 \\
S_1: e_5 \rightarrow e_5
\end{array}
\quad \Leftrightarrow \quad S_1 = \begin{bmatrix}
1 & 2\delta & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

This choice fails to satisfy our requirements, as

\[
S_1^{-1}AS_1 = \begin{bmatrix}
0 & 1 & 0 & 0 & \delta \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\quad \text{and} \quad
S_1^*HS_1 = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 1 & 2\delta & 0 & \delta \\
0 & 0 & 0 & 0 & 1 \\
0 & \delta & 0 & 1 & 0
\end{bmatrix},
\]

have nonzero off diagonal blocks. The resolution is to choose $S_1$ to map not only the vectors of the first chain of $A_0$ to that of $A$, but also map the other chains appropriately. Define

\[
g_4 = \begin{bmatrix}
0 & 0 & 0 & 1 & 0
\end{bmatrix}^T, \quad g_5 = \begin{bmatrix}
0 & -\delta & 0 & \delta^2 & 1
\end{bmatrix}^T.
\]

It is straightforward to check that $[x, y]_H = 0$ for $x \in \text{span}\{g_1, g_2, g_3\}$ and $y \in \text{span}\{g_4, g_5\}$; that is, the first root subspace $\text{span}\{g_1, g_2, g_3\}$ is $H$-orthogonal to all other root subspaces, in this case $\text{span}\{g_4, g_5\}$.

Choosing the matrix $S_1$ such that

\[
\begin{array}{c}
S_1: e_1 \rightarrow g_1 \\
S_1: e_2 \rightarrow g_2 \\
S_1: e_3 \rightarrow g_3 \\
S_1: e_4 \rightarrow g_4 \\
S_1: e_5 \rightarrow g_5
\end{array}
\quad \Leftrightarrow \quad S_1 = \begin{bmatrix}
1 & 2\delta & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & -\delta \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & \delta^2 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix},
\]

as desired.
Lipschitz stability of canonical Jordan bases of $H$-selfadjoint matrices

results in

$$S_1^{-1}AS_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad S_1^*HS_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 2\delta & 0 \\ 1 & 2\delta & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and so

$$(A, H) \xrightarrow{S_1} (A_1, H_1),$$

and since all perturbations are of the order $\delta$, it is easy to see the required bound is satisfied. Hence $S_1$ is a correct choice in that it decouples and allows us to proceed inductively.

**7.1.2. Second step. Zeroing sub–antidiagonal entries of $H_1$. Constructing $S_2$.** In the second step we produce a matrix $S_2$ that eliminates the $2\delta$ sub–antidiagonal elements of the upper left submatrix of $H_1$ to produce the desired structure. Define

$$S_2 = \begin{bmatrix} 1 & -\frac{3\delta^2 - \delta}{2} & 0 & 0 \\ 0 & 1 & -\delta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

so that

$$S_2^{-1}S_1^{-1}AS_1S_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad S_2^*S_1^*HS_1S_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 + \delta^2 \end{bmatrix}.$$  

**7.1.3. Third step. Combining similitude matrices $S_1$ and $S_2$.** Define the matrix $S = S_1S_2$, and so

$$S = \begin{bmatrix} 1 & -\frac{3\delta^2 - \delta}{2} & 0 & 0 \\ 0 & 1 & -\delta & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

can easily be seen to satisfy the bound as all differences from identity are on the order of $\delta$. The previous steps have shown that

$$(A, H) \xrightarrow{S_1} (A_1, H_1') \xrightarrow{S_2} (A_1, H_1)$$

and so

$$(A, H) \xrightarrow{S} (A_1, H_1)$$

of the desired form. With the first block decoupled, we can proceed inductively.

**7.2. Proof of Lemma 5.2**

Part (i) follows from Lemma 6.14. Let us prove part (ii). Following the structure of the above example we prove Lemma 5.2 in three steps.

**7.2.1. First Step. Mapping $A \rightarrow A_1$. Constructing $S_1$.** In this section we prove the following proposition.

**Proposition 7.1 (First decoupling step).** Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed $H_0$-selfadjoint matrix, both $A_0$ and $H_0$ in the canonical form described in Theorem 2.1, and assume that $A_0$ has a real eigenvalue $\lambda_1$. Denote

$$A_0 = \begin{bmatrix} J(\lambda_1) & 0 \\ 0 & A_0 \end{bmatrix} \quad \text{and} \quad H_0 = \begin{bmatrix} \epsilon_1P_1 & 0 \\ 0 & \tilde{H}_0 \end{bmatrix};$$

i.e. consider the partition that singles out the largest, say, $m \times m$, Jordan block $J(\lambda_1)$ of $A_0$. Then there are positive constants $K$ and $\delta$ (depending on $A_0$ and $H_0$ only) such that for any $H$-selfadjoint matrix $A$ with the same Jordan structure as $A_0$ satisfying

$$\|A - A_0\| + \|H - H_0\| < \delta,$$

descriptions hold.
Proof. (i). A has an eigenvalue $\mu_1$ satisfying
\[ |\lambda_1 - \mu_1| \leq K\|A - A_0\|. \tag{7.2} \]

(ii). There exists a similitude matrix $S_1$, i.e., $(A, H) \xrightarrow{S_1} (A_1, H_1)$ such that matrices $A_1$ and $H_1$ have the form
\[ A_1 = \begin{bmatrix} J(\mu_1) & 0 \\ 0 & \tilde{A}_1 \end{bmatrix} \quad \text{and} \quad H_1 = \begin{bmatrix} \tilde{R}_1 & 0 \\ 0 & \tilde{H}_1 \end{bmatrix}, \tag{7.3} \]

with some $\mu_1, \tilde{A}_1, \tilde{R}_1, \tilde{H}_1$, and $S_1$ satisfies the bound
\[ \|I - S_1\| \leq K(\|A - A_0\| + \|H - H_0\|). \tag{7.4} \]

Procedure to construct the matrix $S_1$ of the Proposition 7.1

- **Defining $f_k$'s, $F_2$ and $F_2$.** Since $A_0$ of Proposition 7.1 has canonical Jordan form, we can set $f_k = e_k$, the standard basis vectors. Further, set
\[ F_1 = \text{span}\{f_1, \ldots, f_m\}, \quad F_2 = \text{span}\{f_{m+1}, \ldots, f_n\}, \]

where $m$ is the size of the largest Jordan block $J(\lambda)$ of $A_0$ in (7.1).

- **Defining $g_k$'s, $G_1$ and $G_2$.** Let
\[ G_1 = \text{span}\{g_1, \ldots, g_m\}, \]

where $\{g_k\}_{k=1}^m$ are those described in Lemma 6.14 and satisfying (6.34). Now, in order to define the rest of the vectors $\{g_k\}_{k=m+1}^n$ we need to make several observations.

- First, $F_1 \perp F_2$. This follows easily from the block-diagonal structure of $H_0$ in (7.1) and the definition of indefinite inner product.

- Second, by Lemma 6.8 and (6.34) we have $\theta(F_1, G_1) \leq K_0\|A - A_0\|$, where the constant $K_0 > 0$ depends on $A_0$ only and can be fixed in advance. Hence, if we define
\[ G_2 = G_1^\perp, \]

then by Lemma 6.10 we will have
\[ \theta(F_2, G_2) \leq K_1(\|A - A_0\| + \|H - H_0\|). \]

Therefore, if for $k = m + 1, \ldots, n$ we define $g_k = P_{G_2}f_k$ then
\[ \|f_k - g_k\| \leq K(\|A - A_0\| + \|H - H_0\|). \tag{7.5} \]

Again, the analysis shows that the constant $K_1 > 0$ depends on $A_0$ and $H_0$ only and can be fixed in advance.

- **Defining $S_1$.** Let us define $S_1 : f_k \to g_k$ for $k = 1, \ldots, n$.

We are now ready to complete the proof of (ii). It is easy to see that the bounds (6.34) and (7.5) imply the desired bound (7.4), the proof of it follows along the lines of the proof of Proposition 4.2, literally. Hence it remains only to show that $S_1$ yields via $(A, H) \xrightarrow{S_1} (A_1, H_1)$ matrices $A_1$ and $H_1$ having the structure shown in (7.3).

First, recall that $S_1 : \mathcal{F}_k \to \mathcal{G}_k$. This fact and $H_1 = S_1^*HS$ allow us to deduce
\[ \mathcal{F}_1 \perp H_1 \mathcal{F}_2 \tag{7.6} \]

from $G_1 \perp H G_2$. Relation (7.6) means that in the standard basis $\{f_k\}$ the matrix $H_1$ has the block diagonal form shown in (7.3).

Secondly, $S_1 : f_k \to g_k$ means that $S_1$ maps the Jordan chain $\{f_k\}_{k=1}^n$ of $A_0$ to a Jordan chain $\{g_k\}_{k=1}^n$ of $A_1 = S_1^{-1}AS_1$. Hence $\mathcal{F}_2$ is $A_1$-invariant, and moreover, the first block column of $A_1$ must have the form shown in (7.3). Finally, it is well-known that an $H_1$-orthogonal complement $\mathcal{F}_2$ of an
Remark 7.2 (First step in the case when \( A_0 \) has only complex eigenvalues). In this case one has to consider the partitioning
\[
A_0 = \begin{bmatrix} J_k(\lambda) & 0 \\ 0 & J_k(\bar{\lambda}) \end{bmatrix}, \quad H_0 = \bar{I},
\]
and the proof follows the same lines as above (as it was in Section 4.3).

7.2.2. Second Step. Zeroing sub–antidiagonal entries of \( H_1 \). Constructing \( S_2 \). In the first step above we have constructed \( S_1 \) and mapped \((A, H) \xrightarrow{S_1} (A_1, H_1)\) shown in (7.3). Here we proceed with the upper left blocks \((J(\mu), \bar{R}_1)\) of \((A_1, H_1)\) and construct \( \tilde{S}_2 \) such that \((J(\mu), \bar{R}_1) \xrightarrow{\tilde{S}_2} (J(\mu), \epsilon P_1)\). Theorem 4.1 implies that such \( \tilde{S}_2 \) exists and it must satisfy
\[
\|I - \tilde{S}_2\| \leq K\|\bar{R}_1 - \epsilon P_1\|.
\]
The latter relation yields
\[
\|I - S_2\| \leq K\|H_1 - H_0\| \tag{7.7}
\]
where \( S_2 = \begin{bmatrix} \tilde{S}_2 & 0 \\ 0 & I \end{bmatrix} \) satisfy \((A_1, H_1) \xrightarrow{S_2} (A_2, H_2)\).

7.2.3. Third Step. Combining the similitude matrices \( S_1 \) and \( S_2 \). In the two previous steps we have constructed matrices \( S_1 \) and \( S_2 \) such that
\[
(A, H) \xrightarrow{S_1} (A_1, H_1) \xrightarrow{S_2} (A_2, H_2).
\]
Hence \( S = S_1S_2 \) satisfies \((A, H) \xrightarrow{S} (A_2, H_2)\) and the desired bound (5.5) now follows from (7.4), (7.7), and Theorem 4.9.

This completes the proof of Lemma 5.2.


The last result to prove in the flow chart of Section 2.5 is Theorem 2.6. We will prove it in Section 8.1. As a corollary, we will then deduce in Section 8.2 a classical result of [GLR83] on the stability of sign characteristic. Both results will be needed in Section 8.3 to define the concept of strong similitude matrix.

8.1. Proof of Theorem 2.6

Proof of Theorem 2.6. First, by the definition of a canonical Jordan basis, the (fixed) vectors
\[
\{ \{ f_r^{(k,s)} \}_{r=0}^{m_k(A_0, \lambda_s) - 1} \}_{s=1, k=1}^{\beta, k=\dim\Ker(A_0 - \lambda_s I)}
\]
are the columns of a certain (fixed) similarity-for-pairs matrix \( F \) (i.e., satisfying
\[
(A_0, H_0) \xrightarrow{F} (J_0, P_0) \tag{8.1}
\]
where \((J_0, P_0)\) is canonical). Indeed, it follows from
\[
F^{-1}A_0F = J_0.
\]
Secondly, \( A \) and \( A_0 \) have the same Jordan structure and hence by Theorem 1.8 there exists \( S \) satisfying (1.7) and such that
\[
(A, H) \xrightarrow{S} (A_1, H_0) \tag{8.2}
\]
where the matrix
\[
A_1 = S^{-1}AS
\]
has the same Jordan bases as \( A_0 \). So,
\[
(A_1, H_0) \xrightarrow{F} (J_1, P_0) \tag{8.3}
\]
where \( J_1 \) is Jordan. In particular, we have that
\[
F^{-1}A_1F = J_1.
\]
It follows that
\[ F^{-1} S^{-1} A SF = J_1. \]
Denoting by
\[ \{ \{ f_r^{(k,s)} \}_{s=0}^{m_k(A,\mu_s)-1} \}_{r=0}^{k=\dim \ker(A-\mu_s I)} \]
the columns of the matrix
\[ G := SF \]
we see that both bases \( \{ f_r^{(k,s)} \} \) and \( \{ g_r^{(k,s)} \} \) are canonical, and the desired (2.8) is, in fact, a reformulation of (1.7).

To introduce in Section 8.5 the concept of (strong) similitude matrix (as opposed to weak similitude matrix of Definition 1.7) we will need the following observation.

**Remark 8.1.** Let \( S \) be a (weak) similitude matrix constructed in the process of the proof of Theorem 1.8. Then we have the following two observations.

(i). Formula (8.4) implies the following useful property of \( S \):
\[ g_r^{(k,s)} = S f_r^{(k,s)} \]
for all \( k, s, r \) within their ranges.

(ii). Using the above notations, let the chains \( \{ f_r^{(k,1)} \} \) correspond to the eigenvalue \( \lambda_1 \) of \( A_0 \), and let the chains \( \{ g_r^{(k,1)} \} \) correspond to the eigenvalue \( \mu_1 \) of \( A \). The flow chart for the proof of Theorem 1.8 in Section 2.5 indicates that Lemma 5.2 was crucial in constructing \( S \). Therefore, by inspecting (5.3) (and also (7.2)) we see that the eigenvalue \( \mu_1 \) of \( A \) had been chosen to be a small perturbation,
\[ |\lambda_1 - \mu_1| \leq K \| A - A_0 \|, \]
of the eigenvalue \( \lambda_1 \) of \( A_0 \).

The concept of strong similitude matrix will be introduced in Section 8.3 using two observations. One of them is the above Remark 8.1, and the second one is one classical result [GLR83] on the stability of sign characteristic recalled next.

### 8.2. Stability of sign characteristic as a consequence of the stability of canonical bases

We begin this subsection with the following simple example.

**Example 8.2 (Computing sign characteristic from a canonical Jordan basis).** Let us return to the example 2.5 and examine one way of computing the sign characteristic of \((J, P)\) from its canonical basis. In this case \( \{e_1, e_2, e_3\} \) and \( \{e_3, e_4\} \) are the two Jordan chains of \( J \), and it is easy to see that
\[ e_1 = 1 = [e_1, e_2]_P, \quad e_2 = -1 = [e_4, e_5]_P. \]

Theorem 2.1 and the argument similar to the one in the above example lead to the following obvious statement.

**Lemma 8.3 (Computing the sign characteristic).** Let \( A \in \mathbb{C}^{n \times n} \) be a fixed \( H \)-selfadjoint matrix. Let
\[ \{ \{ f_r^{(k,s)} \}_{s=0}^{m_k(A,\lambda_s)-1} \}_{r=0}^{k=\dim \ker(A-\lambda_s I)} \]
and let \( \{ \lambda_1, \ldots, \lambda_n \} \) be all real eigenvalues of \( A \) be a fixed canonical Jordan basis of \((A, H)\). Then the sign characteristic (2.5) satisfies
\[ \epsilon_{k,s} = \mathcal{F}_{\{r\}_{r=0}^{k(s)}}, f_{m_k(A,\lambda_s)-1} H, \quad (k = 1, \ldots, \dim \ker(A - \lambda_s I), s = 1, \ldots, \alpha). \]

It is easy to see that Theorem 2.6, part (ii) of Remark 8.1 and Lemma 8.3 imply the following well-known result [GLR83, GLR05, R06].

**Theorem 8.4 (Global stability of the sign characteristic).** Let \( A_0 \in \mathbb{C}^{n \times n} \) be a fixed \( H_0 \)-selfadjoint matrix, and let \( \{ \lambda_1, \ldots, \lambda_n \} \) denote all distinct real eigenvalues of \( A_0 \). Let \( \gamma > 0 \) be such that every real eigenvalue \( \lambda_k \) of \( A_0 \) is the only eigenvalue in the interval \((\lambda_k - \gamma, \lambda_k + \gamma)\). There exists a constant \( \delta > 0 \) (depending on \( A_0, H_0 \) and \( \gamma \) only) such that the following assertion holds. For any \( H \)-selfadjoint matrix \( A \) such that \( A \) has the same Jordan structure as \( A_0 \) and
\[ \| A - A_0 \| + \| H - H_0 \| < \delta, \]
matrix \( A \) has a unique eigenvalue, say, \( \mu_k \) in the interval \((\lambda_k - \gamma, \lambda_k + \gamma)\) and the sign characteristics of \( \lambda_k \) and \( \mu_k \) coincide (up to a rearrangement of the signs corresponding to the same block sizes).
The above result is global, i.e., it assumes that the Jordan structure of $A_0$ is preserved for all eigenvalues. In Corollary 9.6 of Section 9 we will also obtain a local version of this stability result.

8.3. Similitude matrix revisited. Mapping canonical bases

Definition 1.7 introduced weakly similitude matrices $S$. Moreover, up until this point the term “similitude” was understood in the weak sense. However, it had been just observed in Section 8.2 that the similitude matrix $S$ had been constructed in such a way that it has several additional properties. We therefore define next the similitude matrix as weakly similitude matrix obeying those additional restrictions.

Definition 8.5 (Strong similitude-for-pairs matrix). Let $A_0$ be $H_0$-selfadjoint and $A$ be $H$-selfadjoint. A matrix $S$ is called a strong similitude matrix (or just a similitude matrix) of the quadruple $(A_0, H_0, A, H)$ if there exist two canonical Jordan bases $(f_i^{(k,s)})$ and $(g_i^{(k,s)})$ of $A_0$ and $A$, respectively, such that

- $(8.5)$ holds.
- $S$ in $(8.5)$ maps Jordan chains $(f_i^{(k,s)})$ corresponding to real eigenvalues of $A_0$ to Jordan chains $(g_i^{(k,s)})$ corresponding to real eigenvalues of $A$, and the same property holds for nonreal eigenvalues as well.
- The mapping $S : (f_i^{(k,s)}) \rightarrow (g_i^{(k,s)})$ preserves the sign characteristic for each Jordan chain.

In this case pairs $(A_0, H_0)$ and $(A, H)$ are called (strongly) similitude.

Remark 8.6 (The difference between similitude relation and weak similitude relation). We start with three obvious observations.

- Both weak similitude and strong similitude are equivalence relations.
- The equivalence class of all matrices that are strongly similitude to a given pair $(A, H)$ is a subset of all matrices that are weakly similitude to $(A, H)$.
- Finally, each pair $(A, H)$ is similitude to its canonical form $(J, P)$ described in Theorem 2.1.

Hence it is of interest to describe both weak and strong similitude relations in term of canonical forms. The comparison of Definitions 1.7 and 8.5 yields the following facts.

(i). Two pairs $(A, H)$ and $(B, G)$ (where $A$ and $B$ are $H$-selfadjoint and $G$-selfadjoint, respectively) are weakly similitude if the matrices $J_A$ and $J_B$ (of their canonical forms $(J_A, P_H)$ and $(J_B, P_G)$) have the same Jordan structure, i.e., there is a bijection $J : \{\lambda_1, \ldots, \lambda_3\} \rightarrow \{\mu_1, \ldots, \mu_3\}$ such that $J_B$ is obtained from $J_A$ by replacing $\lambda_k$’s by $\mu_k$’s. Here $\{\lambda_1, \ldots, \lambda_3\}$ and $\{\mu_1, \ldots, \mu_3\}$ are the sets of all eigenvalues of $A$ and $B$, respectively.

There is no restriction on the structure of sip matrices $P_H$ and $P_G$.

(ii). Two pairs $(A, H)$ and $(B, G)$ are strongly similitude if, in addition to the description of part (i), we also have $P_H = P_G$ which means that not only $A$ and $B$ have same Jordan structure, but they also share the same sign characteristic.

Recall that in Theorem 1.8 we proved the existence of weak similitude matrix $S$. It is now clear that it is the stability of sign characteristic that allowed us to conclude that Theorem 1.8, in fact, claims the existence of a strong similitude matrix $S$.

The “local” Definition 8.5 has, in certain circumstances, some advantages over the “global” Definition 1.7. For instance, in the next section it will be adapted to derive a variant of the main result for the case when the Jordan structure is preserved only for a selection of the eigenvalues.

9. Perturbations partially preserving Jordan structure

9.1. Balanced partitions

In this section, we present an extension of the case considered thus far to the case of perturbations that preserve the Jordan structure only for some selection of the eigenvalues. To be specific, let $A_0$ be $H_0$-selfadjoint, and let

$$\sigma(A_0) = \sigma_1(A_0) \cup \sigma_2(A_0) \quad \text{(with } \sigma_1(A_0) \cap \sigma_2(A_0) = \emptyset)$$

be a partition of the set $\sigma(A_0)$ of all eigenvalues of $A_0$. Recall that in accordance with Remark 2.2, $\sigma(A_0)$ is symmetric with respect to the real axis. In this section we consider only what we suggest to call balanced partitions, i.e., those for which $\sigma_1(A_0)$ is symmetric about the real axis as well (in fact, $\sigma_2(A_0)$ will be automatically symmetric as well in this case).
In the rest of the paper we extend the results of Sections 2 – 8 to the situation where Jordan structure is assumed to be preserved for the eigenvalues in $\sigma_1(A_0)$ only.

9.2. Basic definitions. $\sigma_1$–partial Jordan structure. $\sigma_1$–partial similitude matrix

In this section we provide a number of counterparts for the basic definitions and facts of Sections 1 and 2.

Definition 9.1 (A counterpart of Definition 1.2).

- (Same $\sigma_1$–partial Jordan structure). Let (9.1) be a balanced partition of the set of all eigenvalues of $A_0$. Matrices $A_0$ and $A$ are said to have the same $\sigma_1$–partial Jordan structure if there is a balanced partition

$$\sigma(A) = \sigma_1(A) \cup \sigma_2(A)$$

(9.2)

and a bijection $f: \sigma_1(A_0) \to \sigma_1(A)$ such that if $\mu = f(\lambda)$, then $\lambda$ and $\mu$ have the same Jordan block sizes.

- (Same $\sigma_1$–partial Jordan bases). In this case, matrices $A_0$ and $A$ are said to have the same $\sigma_1$–partial Jordan bases if the following statement is true. If $\mu = f(\lambda)$, then every Jordan chain of $A_0$ corresponding to $\lambda$ is also a Jordan chain of $A$ corresponding to $\mu$ (and automatically vice versa).

The following statement is a counterpart of Remark 1.3

Remark 9.2 (Same $\sigma_1$–partial Jordan bases). Let (9.1) and (9.2) be balanced partitions of $A_0$ and $A$, respectively. The matrices $A_0$ and $A$ have the same $\sigma_1$–partial bases if the following statement holds. If, for an invertible $T$, the matrix $T^{-1}A_0T$ has the form

$$T^{-1}A_0T = \begin{bmatrix} J_0 & \ast \\ 0 & M_0 \end{bmatrix}, \quad \text{with } \sigma(J_0) = \sigma_1(A_0), \sigma(M_0) = \sigma_2(A_0),$$

where $J_0$ is in a canonical Jordan form, then $T^{-1}AT$ has the form

$$T^{-1}AT = \begin{bmatrix} J_1 & \ast \\ 0 & M_1 \end{bmatrix}, \quad \text{with } \sigma(J_1) = \sigma_1(A), \sigma(M_1) = \sigma_2(A),$$

where $J_1$ is also in a canonical Jordan form.

The next definition is a counterpart of Definition 2.4.

Definition 9.3 (Canonical $\sigma_1$–partial Jordan basis. $\sigma_1$–partial sign characteristic). Let $A_0$ be an $H_0$–selfadjoint matrix. Let $\sigma(A_0) = \{\lambda_1, \ldots, \lambda_{2\beta}\}$ be the set of all eigenvalues of $A_0$ and let

$$\{\{f^{(k,s)}_r\}_{r=0}^{m_k(A_0, \lambda_s) - 1} \}_{s=1, k=1}$$

be its canonical Jordan basis. Let (9.1) be a balanced partition with $\sigma_1(A_0) = \{\lambda_{j_1}, \ldots, \lambda_{j_r}\}$.

- The subset

$$\{\{g^{(k,s)}_r\}_{r=0}^{m_k(A_0, \lambda_s) - 1} \}_{s=1, k=1}$$

of (9.3) is called a canonical $\sigma_1$–partial Jordan basis of $A_0$.

- A subset of signs in the sign characteristic of $(A_0, H_0)$ corresponding to its canonical $\sigma_1$–partial Jordan basis is called $\sigma_1$–partial sign characteristic.

In words, a $\sigma_1$–partial Jordan basis of $A_0$ is just a selection of its canonical Jordan chains corresponding to the eigenvalues belonging to the subset $\sigma_1(A_0)$, and the same is true for $\sigma_1$–partial sign characteristic.

We are now ready to formulate a counterpart of Definition 8.5 and to define a “local” version of a similitude matrix.

Definition 9.4 ($\sigma_1$–partial similitude matrix). Let $A_0$ be $H_0$–selfadjoint and $A$ be $H$–selfadjoint. A matrix $S$ is called a $\sigma_1$–partial similitude matrix of the quadruple $(A_0, H_0, A, H)$ if there exist two $\sigma_1$–partial canonical Jordan bases $\{f^{(k,s)}_r\}$ and $\{g^{(k,s)}_r\}$ of $A_0$ and $A$, respectively, such that

- (8.5) holds.

- $S$ in (8.5) maps Jordan chains $\{f^{(k,s)}_r\}$ corresponding to real eigenvalues in $\sigma_1(A_0)$ to Jordan chains $\{g^{(k,s)}_r\}$ corresponding to real eigenvalues in $\sigma_1(A)$, and the same property holds for nonreal eigenvalues in $\sigma_1(A_0)$ and $\sigma_1(A)$ as well.

- For each Jordan chain corresponding to real eigenvalues in $\sigma_1(A_0)$ the mapping $S : \{f^{(k,s)}_r\} \to \{g^{(k,s)}_r\}$ preserves the sign characteristic.

In this case pairs $(A_0, H_0)$ and $(A, H)$ are called $\sigma_1$–partial similitude.
9.3. Lipschitz stability of $\sigma_1$-partial similitude matrices

The main result of this section, an extension of Theorem 2.6 (and hence of Theorem 1.8), is stated next.

**Theorem 9.5 (Lipschitz stability for $\sigma_1$-partial perturbations).** Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed $H_0$-selfadjoint matrix, and let (9.1) be a balanced partition for $A_0$ with $\sigma_1(A_0) = \{\lambda_1, \ldots, \lambda_r\}$. Let
\[
\{(f^{(k,s)}_r)_{r=0}^{m_k(A_0, \lambda_s)}\}_{s=1, k=1}^{1, m_k(A_0, \lambda_s)} = \text{dim Ker}(A - \lambda_s I)
\]
be a $\sigma_1$-partial canonical Jordan basis of $A_0$. Finally, let $\Gamma$ be a simple, closed rectifiable contour such that $A_0$ does not have eigenvalues on $\Gamma$ and $\sigma_1(A_0)$ is the set of all eigenvalues of $A_0$ inside $\Gamma$.

There exist constants $K, \delta > 0$ (depending on $A_0$ and $H_0$ only) such that for any $H$-selfadjoint matrix $A$ having the same $\sigma_1$-partial Jordan structure as $A_0$ (where $\sigma_1(A)$ is the set of all eigenvalues of $A$ inside $\Gamma$) and satisfying
\[\|A - A_0\| + \|H - H_0\| < \delta,\]
the following assertion holds.

(i). **(Lipschitz stability of $\sigma_1$-partial similitude matrices)** The pairs $(A_0, H_0)$ and $(A, H)$ are $\sigma_1$-partially similitude, and there exists a $\sigma_1$-partial similitude matrix $S$ satisfying
\[\|S - I\| \leq K \left(\|A - A_0\| + \|H - H_0\|\right).\] (9.5)

(ii). **(Lipschitz stability of $\sigma_1$-partial canonical Jordan bases)** There exists a $\sigma_1$-partial canonical Jordan basis $\{\{g^{(k,s)}_r\}_{r=0}^{m_k(A, \mu_s)}\}_{s=1, k=1}^{1, m_k(A, \mu_s)}$ of $A$ such that
\[\|g^{(k,s)}_r - f^{(k,s)}_r\| \leq K \left(\|A - A_0\| + \|H - H_0\|\right)\] (9.6)
for $k = 1, \ldots, \gamma$ and all $k, r$ within their ranges.

Before proving the above theorem in Section 9.4 we formulate the following obvious corollary.

**Corollary 9.6 (Stability of $\sigma_1$-partial sign characteristic).** Let $A_0 \in \mathbb{C}^{n \times n}$ be a fixed $H_0$-selfadjoint matrix, and let (9.1) be a balanced partition for $A_0$. There exist a constant $\delta > 0$ (depending on $A_0$ and $H_0$ only) such that for any $H$-selfadjoint matrix $A$ having the same $\sigma_1$-partial Jordan structure as $A_0$ and satisfying
\[\|A - A_0\| + \|H - H_0\| < \delta,\]
the $\sigma_1$-partial sign characteristics of $A$ and $A_0$ coincide.

9.4. Proof of Theorem 9.5

Actually, the proof of Theorem 9.5 can be derived by just adapting all results of Sections 2 – 8 to the $\sigma_1$-partial case. For example, here is a counterpart of the Theorem 3.1

Thus to replace $H_0$ and $G_0$ of (3.1) by $H_1$ and $G_1$, respectively. The point of this replacement is that if we partition the canonical pair $(A_0, H_0)$ as
\[A_0 = \begin{bmatrix} A_1^{(0)} & 0 \\ 0 & A_2^{(0)} \end{bmatrix}, \quad H_0 = \begin{bmatrix} H_1^{(0)} & 0 \\ 0 & H_2^{(0)} \end{bmatrix}, \quad \text{with } \sigma(A_1^{(0)}) = \sigma_1(A_0), \sigma(A_2^{(0)}) = \sigma_2(A_0),\]
then the pair $(A_1, H_1)$ should look like
\[A_1 = \begin{bmatrix} A_1^{(1)} & 0 \\ 0 & A_2^{(1)} \end{bmatrix}, \quad H_1 = \begin{bmatrix} H_1^{(1)} & 0 \\ 0 & H_2^{(1)} \end{bmatrix}, \quad \text{with } \sigma(A_1^{(1)}) = \sigma_1(A), \sigma(A_2^{(1)}) = \sigma_2(A),\]

**Theorem 9.7 (Reduction to the canonical form).** Suppose the result of Theorem 9.5 is true for each pair $(A_0, H_0)$ in the canonical form as defined in the Definition 2.3. Then the result of Theorem 9.5 is true for all pairs $(B_0, G_0)$, where $B_0$ is $G$-selfadjoint.

**Sketch of the proof.** The proof follows the lines of the proof of Theorem 3.1 with only minor modifications. We therefore highlight only the major differences. Specifically, one has to modify the diagram (3.1) as follows.

\[\begin{array}{c}
(A, H) \xrightarrow{S} (A_1, H_1) \\
\uparrow \quad \uparrow \\
(B, G) \xrightarrow{R=TS^{-1}} (B_1, G_1)
\end{array}\] (9.7)

Thus to replace $H_0$ and $G_0$ of (3.1) by $H_1$ and $G_1$, respectively. The point of this replacement is that if we partition the canonical pair $(A_0, H_0)$ as
\[A_0 = \begin{bmatrix} A_1^{(0)} & 0 \\ 0 & A_2^{(0)} \end{bmatrix}, \quad H_0 = \begin{bmatrix} H_1^{(0)} & 0 \\ 0 & H_2^{(0)} \end{bmatrix}, \quad \text{with } \sigma(A_1^{(0)}) = \sigma_1(A_0), \sigma(A_2^{(0)}) = \sigma_2(A_0),\]
then the pair $(A_1, H_1)$ should look like
\[A_1 = \begin{bmatrix} A_1^{(1)} & 0 \\ 0 & A_2^{(1)} \end{bmatrix}, \quad H_1 = \begin{bmatrix} H_1^{(1)} & 0 \\ 0 & H_2^{(1)} \end{bmatrix}, \quad \text{with } \sigma(A_1^{(1)}) = \sigma_1(A), \sigma(A_2^{(1)}) = \sigma_2(A),\]
where the pair \((A_1^{(1)}, H_1^{(1)})\) is canonical, i.e., \(A_1^{(1)}\) is Jordan and \(H_1^{(1)}\) is sip. Considering matrices \((A_1, H_1)\) that have canonical structure only in their leading blocks is exactly what is needed for proving the desired result in the \(\sigma_1\)-partial case.

The rest of the arguments follow the lines of the proof of Theorem 3.1.

Using the above result it is possible to complete the proof of Theorem 9.5 by adapting the rest of the proof of Theorem 5.1. Indeed, we proved Theorem 5.1 recursively, i.e., by decoupling Jordan blocks one by one. Adaptation of that proof to the \(\sigma_1\)-partial case simply means stopping the decoupling process earlier, after processing all the eigenvalues in \(\sigma_1\).

However, instead of asking the reader to inspect the proof of Theorem 5.1 in Sections 4, 5, and 7 (see, e.g., the flow chart in Section 2.5) we prefer to give a short direct proof.

**Proof of Theorem 9.5.** The proof of (i) is the main part of the proof, the claim (ii) is just a corollary of (i). (i). Let \(\sigma_1(A_0) = \{\lambda_1, \ldots, \lambda_n\}\), and let the canonical pair \((A_0, H_0)\) be partitioned as

\[
A_0 = \begin{bmatrix}
J^{(0)}(\lambda_1) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & J^{(0)}(\lambda_i) & 0 \\
0 & \cdots & 0 & A_0^{(0)}
\end{bmatrix},
H_0 = \begin{bmatrix}
P^{(0)}_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & P^{(0)}_\gamma & 0 \\
0 & \cdots & 0 & H^{(0)}_{\gamma+1}
\end{bmatrix},
\]

where each pair \((J^{(0)}(\lambda_i), P^{(0)}_i)\) is in the canonical form, \(\sigma(J^{(0)}(\lambda_i)) = \{\lambda_i\}\) and \(\sigma(A_0^{(0)}) \subset \sigma(A_0)\).

Further, let \(\Gamma_i\) be a set of small non-intersecting circles such that \(\lambda_i\) is the only eigenvalue of \(A_0\) inside each circle \(\Gamma_i\). Since \(A\) has the same \(\sigma_1\)-partial Jordan structure as \(A_0\) there is \(\delta > 0\) guaranteeing that \(A\) has only one eigenvalue, say, \(\mu_i\) inside each \(\Gamma_i\), i.e., \(\sigma_1(A) = \{\mu_1, \ldots, \mu_\gamma\}\). Denote

\[
M_i = \Re(A_0, \lambda_i), \quad N_i = \Re(A, \mu_i), \quad i = 1, \ldots, \gamma,
\]

and

\[
M_{\gamma+1} = (M_1 + \ldots + M_\gamma)^{[1]|\lambda|}, \quad N_{\gamma+1} = (N_1 + \ldots + N_\gamma)^{[1]|\mu|};
\]

Using Lemmas 6.10, 6.11, and 6.12 one can construct \(S_1\) such that

\[
S_1M_i = N_i, \quad i = 1, \ldots, \gamma + 1
\]

and

\[
\|I - S_1\| \leq K_1(\|A - A_0\| + \|H - H_0\|)
\]

for some \(K_1\) depending on \(A_0\) and \(H_0\) only. Define \((A_1, H_1)\) by

\[
(A, H) S_1 \rightarrow (A_1, H_1).
\]

Clearly, (9.8) and the fact that \(M_i\) are \(H_0\)-orthogonal and \(N_i\) are \(H\)-orthogonal imply that the matrices in \((A_1, H_1)\) have the same block form

\[
A_1 = \begin{bmatrix}
A^{(1)}(\mu_1) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & A^{(1)}(\mu_\gamma) & 0 \\
0 & \cdots & 0 & A^{(1)}_{\gamma+1}
\end{bmatrix},
H_1 = \begin{bmatrix}
H^{(1)}_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & H^{(1)}_{\gamma} & 0 \\
0 & \cdots & 0 & H^{(1)}_{\gamma+1}
\end{bmatrix},
\]

as \((A, H)\). Indeed, observe that at the moment the pairs \((A^{(1)}(\mu_i), H^{(1)}_i)\) are not in the canonical form yet. Using Theorem 5.1, one finds matrices \(S_i^{(2)}\) such that

\[
(A^{(1)}(\mu_i), H^{(1)}_i) S_i^{(2)} \rightarrow (J^{(1)}(\mu_i), P^{(1)}_i), \quad i = 1, 2, \ldots, \gamma,
\]
where the pairs $(J^{(1)}(\mu_i), P_i^{(1)})$ are canonical. Defining
\[
S_2 = \begin{bmatrix}
S_1^{(2)} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & S_\gamma^{(2)} & 0 \\
0 & \cdots & 0 & I
\end{bmatrix}
\]
one sees that
\[
(A_1, H_1) \overset{S_2}{\longrightarrow} \left( \begin{bmatrix}
J^{(1)}(\lambda_1) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & J^{(1)}(\lambda_\gamma) & 0 \\
0 & \cdots & 0 & A_{\gamma+1}^{(1)}
\end{bmatrix}, \begin{bmatrix}
P_1^{(0)} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & P_\gamma^{(0)} & 0 \\
0 & \cdots & 0 & H_{\gamma+1}^{(1)}
\end{bmatrix} \right)
\]
and by (5.1) we have
\[
\|I - S_2\| \leq K_2(\|A_1 - A_0\| + \|H_1 - H_0\|).
\]
Clearly, $S = S_1S_1$ is the desired $\sigma_1$–partial similitude matrix. Finally, combining (9.9) and (9.10) and using Lemma 4.8 we obtain the bound (9.5).

(ii). The bound (9.5) and the relation $Sf^{(k,s)} = g^{(k,s)}$ yields the desired bound (9.6).

\[\square\]

Acknowledgments

The authors would like to express their sincere appreciation and gratitude for the many helpful suggestions provided by Michael Neumann during the entire process of working on this manuscript. The quality of the paper has been improved as a result.

References


T. Bella* and V. Olshevsky* and U. Prasad*