UNIVERSITÄT BREMEN

Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices

Peter Benner **Volker Mehrmann**
Hongguo Xu

Report 00–01

Berichte aus der Technomathematik

Report 00–01 **Januar 2000**

Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices *

Peter Benner[†] Volker Mehrmann[‡] Hongguo Xu[§]

January 27, 2000

Abstract

We study the perturbation theory for the eigenvalue problem of a formal matrix product $A_1^{s_1} \cdots A_p^{s_p}$, where all A_k are square and $s_k \in \{-1, 1\}$. We generalize the classical perturbation results for matrices and matrix pencils to perturbation results for generalized deflating subspaces and eigenvalues of such formal matrix products. As an application we then extend the structured perturbation theory for the eigenvalue problem of Hamiltonian matrices to Hamiltonian/skew-Hamiltonian pencils.

Keywords. Perturbation theory, eigenvalue problem, formal matrix products, periodic QZ algorithm, Hamiltonian matrix, skew-Hamiltonian matrix, deflating subspace.
AMS subject classification. 65F15, 93B40, 93B60, 65H17.

1 Introduction

The perturbation theory for eigenvalues, eigenvectors and deflating subspaces of matrices and matrix pencils is well established, see the monograph [24] for the classical theory and further references. In this paper we extend some of these results to formal matrix products $A_1^{s_1} \cdots A_p^{s_p}$ for a given set of p square matrices $A_1, \ldots, A_p \in \mathbb{C}^{n \times n}$ and p parameters $s_1, \ldots, s_p \in \{-1, 1\}$. Here if $s_j = -1$ the inverse of the matrix A_j is not required to exist but the inverse is considered only formally to simplify the notation. Our interest in such matrix products arises from applications in the computation of deflating subspaces of Hamiltonian/skew-Hamiltonian pencils, see [2, 3] and from the computation of the *periodic Schur decomposition* introduced in [9, 16]. Other applications of such formal products of matrices are monodromy relations arising for instance in discrete-time periodic (descriptor) systems [1, 8, 19, 25].

For $A_1^{s_1} \cdots A_p^{s_p}$ and $s_1, \ldots, s_p \in \{-1, 1\}$ as described, it is known [9, 16] that there exist p unitary matrices $Q_1, \ldots, Q_p \in \mathbb{C}^{n \times n}$ such that for $Q_{p+1} := Q_1$ and

$$q_k = \frac{1 - s_k}{2}, \quad k = 1, \ldots, p \tag{1}$$

*Supported by *Deutsche Forschungsgemeinschaft* within Project A4 of Sonderforschungsbereich 393, Numerische Simulation auf massiv parallelen Rechnern

[†]Zentrum für Technomathematik, Fachbereich 3/Mathematik und Informatik, Universität Bremen, D-28334 Bremen, Germany; benner@math.uni-bremen.de.

[‡]Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany; mehrmann@mathematik.tu-chemnitz.de

[§]Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106-7058, USA; hxx7@po.cwru.edu

all the matrices

$$R_k = Q^H_{k+q_k} A_k Q_{k+1-q_k} = \begin{cases} Q^H_k A_k Q_{k+1} & s_k = 1 \\ Q^H_{k+1} A_k Q_k & s_k = -1 \end{cases} \quad (2)$$

are upper triangular for $k = 1, \ldots, p$. Such a form is called *periodic Schur form of a formal product.*

The periodic Schur form is the generalization of the usual Schur form for a square matrix A or the generalized Schur form for a square matrix pencil $A - \lambda B$, which are the special cases with $p = 1$, $s_1 = 1$, and $p = 2$ and $s_1 = 1, s_2 = -1$ or $s_1 = -1, s_2 = 1$, respectively.

Numerical methods for computing the periodic Schur decomposition (2) were introduced in [9, 16]. These methods, the *periodic QR algorithm and QZ algorithm* are direct generalizations of the QR and QZ algorithm, e.g., [13, 15, 22, 26].

If all the matrices A_k corresponding to $s_k = -1$ are nonsingular, then for

$$B_1 = A_1^{s_1} \cdots A_p^{s_p}, \quad \ldots, \quad B_k = A_k^{s_k} \cdots A_p^{s_p} A_1^{s_1} \cdots A_{k-1}^{s_{k-1}}, \quad \ldots, \quad B_p = A_p^{s_p} A_1^{s_1} \cdots A_{p-1}^{s_{p-1}}, \quad (3)$$

the periodic Schur form (2) simultaneously gives the Schur forms of $B_1, \ldots, B_p$. In fact from (2) we have $R_k^{s_k} = Q_k^H A_k^{s_k} Q_{k+1}$, which leads to

$$Q_k^H B_k Q_k = R_k^{s_k} \cdots R_p^{s_p} R_1^{s_1} \cdots R_{k-1}^{s_{k-1}}, \quad (4)$$

for $k = 1, \ldots, p$. Observe that in this case all matrices B_k are similar and hence have equal spectra.

It follows that the periodic Schur form is related to the eigenvalue problem for the matrices $B_1, \ldots, B_p$. But the periodic Schur form is more general, since it always exists, regardless of the singularity of the matrices A_k.

In theory, if all the matrices with negative exponent are nonsingular, then the solution of the eigenvalue problem for B_k can be obtained by the QR algorithm [15] applied to the explicitly formed product B_k. However, it is well-known that by forming the product the conditioning of the eigenvalue problem may be increased drastically and furthermore rounding errors, ill-conditioned inverses and subtractive cancellation may lead to a computed product matrix B_k which is nowhere close to the exact formal product. Another problem is that if all Schur forms of B_k are needed, explicitly updating all B_k may be very expensive. For this reason in [9, 16] the periodic QR algorithm was suggested that allows to compute eigenvalues and invariant subspaces of B_k simultaneously without forming the product. Algorithms to compute the products B_k without forming inverses were introduced in [1].

In this paper we discuss the perturbation analysis of the eigenvalue problem for the formal products B_k based on perturbations in the separate factors. The analysis can be viewed as generalization of the usual perturbation theory for eigenvalue problems, see e.g., [24].

We consider the formal product as a map acting on matrix tuples $\mathbf{A} = (A_1, \ldots, A_p)$ in the linear space $\underbrace{\mathbb{C}^{n\times n} \times \ldots \times \mathbb{C}^{n\times n}}_{p}$. The signs s_j are combined in a *sign tuple* $s := (s_1, \ldots, s_p)$.

The connection between the matrix tuples $(A_1, \ldots, A_p)$ and $(B_1, \ldots, B_p)$ in (3) allows to define the eigenstructure corresponding to $\mathbf{A}$. Let $\mathbf{A}$ have a periodic Schur form (2). Let the diagonal elements of R_k be $r_{11;k}, \ldots, r_{nn;k}$ for $k = 1, \ldots, p$.

For an integer $j \in \{1, \ldots, n\}$, if all $r_{jj;k}$ corresponding to $s_k = -1$ are nonzero then $\lambda_j := r_{jj;1}^{s_1} \cdots r_{jj;p}^{s_p}$ is a *finite eigenvalue* of $\mathbf{A}$ associated with the sign tuple s.

If all $r_{jj;k}$ corresponding to $s_k = 1$ are nonzero and some $r_{jj;k}$ corresponding to $s_k = -1$ is zero then $\mathbf{A}$ has an *infinite eigenvalue* $\lambda_j := \infty$.

The spectrum of $\mathbf{A}$, i.e., the set of eigenvalues of B_k including the infinite eigenvalue is denoted by $\Lambda(\mathbf{A})$. If $\mathbf{A}$ has exactly n eigenvalues (including the infinite eigenvalue) then we call $\mathbf{A}$ *regular*. In this paper we will discuss only regular tuples $\mathbf{A}$.

Let nonzero vectors $u_1, \ldots, u_p$ and scalars $\alpha_1, \ldots, \alpha_p$ satisfy

$$A_k u_{k+1-q_k} = \alpha_k u_{k+q_k}, \quad k = 1, \ldots, p, \tag{5}$$

with $u_{p+1} = u_1$. Consider unitary matrices Q_k, $k = 1, \ldots, p$, such that $Q_k e_1 = \frac{1}{\tau_k} u_k$, where $\tau_k = \sqrt{u_k^H u_k}$ and e_1 is the first unit vector. Then we obtain from (5) that

$$Q_{k+q_k}^H A_k Q_{k+1-q_k} = \begin{bmatrix} \frac{\alpha_k \tau_{k+q_k}}{\tau_{k+1-q_k}} & a_k^H \\ 0 & \tilde{A}_k \end{bmatrix}, \quad k = 1, \ldots, p,$$

with index q_k as in (1). If for all s_k with $s_k = -1$ we have $\alpha_k \neq 0$, then

$$\lambda := \prod_{k=1}^{p} \left(\frac{\alpha_k \tau_{k+q_k}}{\tau_{k+1-q_k}}\right)^{s_k} = \prod_{k=1}^{p} \left(\frac{\tau_k}{\tau_{k+1}}\right) \alpha_k^{s_k} = \prod_{k=1}^{p} \alpha_k^{s_k}$$

is a finite eigenvalue of $\mathbf{A}$. Moreover, if for all s_k with $s_k = 1$ we have $\alpha_k \neq 0$ and there exists some k with $s_k = -1$ and $\alpha_k = 0$, then $1/\lambda = 0$ and λ is an infinite eigenvalue. In this sense we call a vector $\mathbf{u} = (u_1, \ldots, u_p)$ with $u_k \neq 0$ for $k = 1, \ldots, p$ a *right eigenvector* of $\mathbf{A}$ corresponding to the eigenvalue λ. As we will see in Section 2 the restriction that $u_k \neq 0$ identifies the eigenvector. If vectors $u_k = 0$ are allowed, then some A_k with $s_k = -1$ is singular and there exist many vectors $\mathbf{u}$ satisfying (5). This is the main difference between the usual eigenvalue problem and that for formal matrix products.

Example 1 *Let $p = 2$, $s_1 = s_2 = 1$ and $A_1 = A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then for $u_1 = u_2 = e_1$,*

$$A_1 u_2 = 0 \cdot u_1, \quad A_2 u_1 = u_2,$$

which implies that (e_1, e_1) is the eigenvector corresponding to the eigenvalue 0. However, if zero vectors are allowed then $u_1 = 0$, $u_2 = e_1$ also satisfy

$$A_2 u_2 = 0 \cdot u_1, \quad A_2 u_1 = 0 \cdot u_2.$$

In order to define deflating subspaces, let $Q_k = [U_k, V_k]$ be a unitary matrix such that

$$Q_{k+q_k}^H A_k Q_{k+1-q_k} = \begin{bmatrix} C_k & F_k \\ 0 & D_k \end{bmatrix} =: T_k, \tag{6}$$

where $C_k \in \mathbb{C}^{m \times m}$ and $U_k \in \mathbb{C}^{n \times m}$ for $k = 1, \ldots, p$. Then

$$A_k U_{k+1-q_k} = U_{k+q_k} C_k, \quad k = 1, \ldots, p,$$

and we call the space spanned by the columns of $\mathbf{U} = (U_1, \ldots, U_p)$ a *right generalized deflating subspace* of $\mathbf{A}$ associated with the sign tuple s corresponding to the spectrum $\Lambda(\mathbf{C})$. Again, if all products B_k, $k = 1, \ldots, p$ in (3) are well defined, then from (4) for each k, the columns of U_k span an orthonormal basis of the invariant subspace of B_k corresponding to the eigenvalues of $C_k^{s_k} \cdots C_p^{s_p} C_1^{s_1} \cdots C_{k-1}^{s_{k-1}}$.

In this paper we derive the perturbation analysis for the eigenvalues and deflating subspaces of formal products $\mathbf{A}$. Some of these results extend the classical perturbation theory for matrices and matrix pencils. We will first study perturbations of the generalized deflating subspace, followed by perturbation results for the eigenvalues. These results will be contained in Section 2. In Section 3 we then study as an application the perturbation theory for Hamiltonian/skew-Hamiltonian pencils under structured perturbations.

We use $\|\cdot\|$ to denote the spectral norm. Throughout this paper we identify k and $k \bmod p$. The smallest singular value of a matrix A is denoted by $\sigma_{\min}(A)$. Finally we denote by $A \otimes B = [a_{ij}B]$ the Kronecker product of matrices A and B and for a matrix $Z = [z_1, \ldots, z_n]$ the operation 'Vec' is defined via $\text{Vec}(Z) = [z_1^T, \ldots, z_n^T]^T$.

2 Perturbation Theory for Generalized Deflating Subspaces and Eigenvalues

In this section we derive the structured perturbation theory for the eigenvalues and generalized deflating subspaces of formal matrix products. We restrict ourselves to the case that the matrix tuple $\mathbf{A}$ is regular. In the case of a nonregular tuple or a tuple that is close to a nonregular tuple, the computation of the generalized deflating subspaces may be an ill-posed problem. Nonregular matrix tuples or tuples close to nonregular tuples already pose a severe difficulty in the case of matrix pencils, see [10, 11, 12, 24].

For the perturbation analysis we will need the following linear transformation. Let $\mathbf{C} = (C_1, \ldots, C_p)$ be a tuple of $m \times m$ matrices with sign tuple $s = (s_1, \ldots, s_p)$ and let $\mathbf{D} = (D_1, \ldots, D_p)$ be another tuple of $l \times l$ matrices with the same sign tuple s. Define a linear transformation Φ on matrix tuples $\mathbf{X} = (X_1, \ldots, X_p) \in \underbrace{\mathbb{C}^{l\times m} \times \ldots \times \mathbb{C}^{l\times m}}_{p}$ via

$$\Phi(\mathbf{X}) = (D_1 X_{2-q_1} - X_{1+q_1} C_1, D_2 X_{3-q_2} - X_{2+q_2} C_2, \ldots, D_p X_{p+1-q_p} - X_{p+q_p} C_p), \qquad (7)$$

with q_k as in (1). In the usual notation for linear operators, Φ is *nonsingular* if $\Phi(\mathbf{X}) = 0$ implies that $\mathbf{X} = 0$, i.e., $X_1 = \ldots = X_p = 0$.

The following result can be viewed as a generalization of the classical existence result for homogeneous Sylvester equations [14]. It is one of the basic tools for the perturbation analysis.

Lemma 1 *For matrix tuples* $\mathbf{C}$ *and* $\mathbf{D}$ *with the same sign tuple* s*, let* Φ *be defined as in (7). Then* Φ *is nonsingular if and only if* $\mathbf{C}$ *and* $\mathbf{D}$ *are regular and* $\Lambda(\mathbf{C}) \cap \Lambda(\mathbf{D}) = \emptyset$.

Proof. Suppose that we have the periodic Schur decompositions

$$\begin{aligned} U_{k+q_k}^H D_k U_{k+1-q_k} &= \tilde{D}_k, \quad k = 1, \ldots, p, \\ V_{k+q_k}^H C_k V_{k+1-q_k} &= \tilde{C}_k, \quad k = 1, \ldots, p, \end{aligned}$$

where all $\tilde{D}_k = [d_{ij;k}]$ are upper triangular and all $\tilde{C}_k = [c_{ij;k}]$ are lower triangular. The latter form can be easily obtained by simultaneously reordering the rows and columns of a periodic Schur form, where all factors are in upper triangular form. Set $\tilde{X}_k = U_k^H X_k V_k$ for $k = 1, \ldots, p$. Then $\Phi(\mathbf{X}) = 0$ if and only if

$$\tilde{\Phi}(\tilde{\mathbf{X}}) := (\tilde{D}_1 \tilde{X}_{2-q_1} - \tilde{X}_{1+q_1} \tilde{C}_1, \ldots, \tilde{D}_p \tilde{X}_{p+1-q_p} - \tilde{X}_{p+q_p} \tilde{C}_p) = 0.$$

Let

$$Z = \begin{bmatrix} G_1 & K_1 & & \\ & \ddots & \ddots & \\ & & \ddots & K_{p-1} \\ K_p & & & G_p \end{bmatrix}, \tag{8}$$

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

$$\begin{array}{lll} G_k = \tilde{C}_k^T \otimes I_l, & K_k = -I_m \otimes \tilde{D}_k, & \text{if } s_k = 1, \\ G_k = I_m \otimes \tilde{D}_k, & K_k = -\tilde{C}_k^T \otimes I_l, & \text{if } s_k = -1 \end{array}$$

and let $x = [\mathrm{Vec}(\tilde{X}_1)^T, \ldots, \mathrm{Vec}(\tilde{X}_p)^T]^T$. Then a simple calculation yields that $\tilde{\Phi}(\mathbf{X}) = 0$ if and only if $Zx = 0$, i.e., Φ is nonsingular if and only if Z is nonsingular. Since all matrices $\tilde{C}_k^T$ and $\tilde{D}_k$ are upper triangular, using the special block structure of Z, a straightforward calculation gives

$$\det Z = \prod_{i=1}^{l} \prod_{j=1}^{m} \left(\prod_{k=1}^{p} \alpha_{ij;k} - \prod_{k=1}^{p} \beta_{ij;k} \right),$$

where

$$\begin{array}{lll} \alpha_{ij;k} = c_{jj;k}, & \beta_{ij;k} = d_{ii;k} & \text{if } s_k = 1, \\ \alpha_{ij;k} = d_{ii;k}, & \beta_{ij;k} = c_{jj;k} & \text{if } s_k = -1. \end{array}$$

Hence $\det Z = 0$ if and only if at least one of the terms $\prod_{k=1}^{p} \alpha_{ij;k} - \prod_{k=1}^{p} \beta_{ij;k}$ is zero. From the definitions of $\alpha_{ij;k}$ and $\beta_{ij;k}$ it is not difficult to see that this is the case if and only if either at least one of the two matrices $\mathbf{C}$, $\mathbf{D}$ is singular or $\Lambda(\mathbf{C}) \cap \Lambda(\mathbf{D}) \neq \emptyset$. □

After these general observations we study perturbations of generalized deflating subspaces.

2.1 Generalized deflating subspaces

Consider a regular matrix tuple $\mathbf{A} = (A_1, \ldots, A_p)$ with sign tuple $s = (s_1, \ldots, s_p)$ and suppose that there exist unitary matrices $Q_k = [U_k, V_k]$ with $U_k \in \mathbb{C}^{n \times m}$ that satisfy (6). The goal of the perturbation analysis is to analyze how much the subspace $\mathrm{range}\, \mathbf{U} := (\mathrm{range}\, U_1, \ldots, \mathrm{range}\, U_p)$ changes if we consider perturbed quantities $A_k + \Delta A_k$, $k = 1, \ldots, p$. In order to get meaningful results, we consider only the case that the generalized deflating subspace is uniquely defined. The following lemma gives a sufficient condition for the uniqueness of the subspace.

Lemma 2 *Consider a regular matrix tuple* $\mathbf{A}$ *with sign tuple* s *having the decomposition (6). If* $\Lambda(\mathbf{C}) \cap \Lambda(\mathbf{D}) = \emptyset$*, then the generalized deflating subspace* $\mathrm{range}\, \mathbf{U}$ *corresponding to* $\Lambda(\mathbf{C})$ *is unique.*

Proof. Suppose there exists another tuple of unitary matrices $\tilde{Q}_k = [\tilde{U}_k, \tilde{V}_k]$ for which (6) also holds, i.e., for $k = 1, \ldots, p$ we have

$$\tilde{Q}_{k+q_k}^H A_k \tilde{Q}_{k+1-q_k} = \begin{bmatrix} \tilde{C}_k & \tilde{F}_k \\ 0 & \tilde{D}_k \end{bmatrix} =: \tilde{T}_k, \tag{9}$$

with $\Lambda(\tilde{\mathbf{C}}) = \Lambda(\mathbf{C})$ and $\Lambda(\tilde{\mathbf{D}}) = \Lambda(\mathbf{D})$. Let $W_k = \tilde{Q}_k^H Q_k =: \begin{bmatrix} W_{11;k} & W_{12;k} \\ W_{21;k} & W_{22;k} \end{bmatrix}$ for $k = 1, \ldots, p$. Then the generalized deflating subspace is unique if and only if the tuple $\mathbf{W}_{21} :=$

$(W_{21;1},\ldots,W_{21;p})$ is the zero tuple. By (6) and (9) we have $\tilde{T}_k W_{k+1-q_k} = W_{k+q_k} T_k$, which implies that $\tilde{D}_k W_{21;k+1-q_k} = W_{21;k+q_k} C_k$ for $k=1,\ldots,p$. Since $\Lambda(\tilde{\mathbf{D}}) \cap \Lambda(\mathbf{C}) = \emptyset$ by employing Lemma 1 we get $\mathbf{W}_{21} = 0$. Hence the generalized deflating subspace is unique. □

Suppose that the matrix tuple $\mathbf{A}$ is perturbed by $\Delta\mathbf{A} := (\Delta A_1,\ldots,\Delta A_p)$ and set

$$\hat{\mathbf{A}} := (\hat{A}_1,\ldots,\hat{A}_p) := (A_1+\Delta A_1,\ldots,A_p+\Delta A_p).$$

We assume that $\mathbf{A}$ is in the form (6), i.e.,

$$T_k = Q^H_{k+q_k} A_k Q_{k+1-q_k} = \begin{bmatrix} C_k & F_k \\ 0 & D_k \end{bmatrix}, \quad k=1,\ldots,p,$$

where $C_k \in \mathbb{C}^{m\times m}$ for $k=1,\ldots,p$. Since the eigenvalues of $\mathbf{C}$ will also be perturbed, we consider an associated perturbed generalized deflating subspace of $\hat{\mathbf{A}}$ corresponding to eigenvalues near to those of $\mathbf{C}$. This subspace is obtained as follows. Introducing

$$\Delta T_k = Q^H_{k+q_k} \Delta A_k Q_{k+1-q_k} =: \begin{bmatrix} \Delta C_k & \Delta F_k \\ E_k & \Delta D_k \end{bmatrix}, \tag{10}$$

we have

$$\hat{T}_k := Q^H_{k+q_k} \hat{A}_k Q_{k+1-q_k} = \begin{bmatrix} C_k+\Delta C_k & F_k+\Delta F_k \\ E_k & D_k+\Delta D_k \end{bmatrix} =: \begin{bmatrix} \hat{C}_k & \hat{F}_k \\ E_k & \hat{D}_k \end{bmatrix}. \tag{11}$$

If $\mathbf{V} := (V_1,\ldots,V_p)$ is an orthonormal basis of a generalized deflating subspace of $\hat{\mathbf{T}} := (\hat{T}_1,\ldots,\hat{T}_p)$, then $(Q_1V_1,\ldots,Q_pV_p)$ is an orthonormal basis of the associated generalized deflating subspace of $\hat{\mathbf{A}}$ corresponding to the same eigenvalues. In the following we therefore consider the perturbation analysis for $\mathbf{T}$ and $\hat{\mathbf{T}}$.

If the perturbations are sufficiently small, then we may simultaneously triangularize the matrices $\hat{T}_1,\ldots,\hat{T}_p$ via unitary matrices of the forms

$$Y_k = \begin{bmatrix} I_m & X_k^H \\ -X_k & I_{n-m} \end{bmatrix} \begin{bmatrix} H_{1k} & 0 \\ 0 & H_{2k} \end{bmatrix}, \tag{12}$$

where $H_{1k} = (I_m + X_k^H X_k)^{-\frac{1}{2}}$ and $H_{2k} = (I_{n-m} + X_k X_k^H)^{-\frac{1}{2}}$ for $k=1,\ldots,p$. Here the matrix $A^{-\frac{1}{2}}$ denotes the Hermitian positive definite square root of an Hermitian positive definite matrix A^{-1}. To make $\hat{\mathbf{T}}$ block upper triangular the matrix tuple $\mathbf{X} := (X_1,\ldots,X_p)$ must solve the system of discrete-time periodic Riccati equations

$$\hat{D}_k X_{k+1-q_k} - X_{k+q_k}\hat{C}_k - E_k + X^H_{k+q_k}\hat{F}_k X_{k+1-q_k} = 0, \quad k=1,\ldots,p. \tag{13}$$

For the analysis of equations of this type see [7, 8]. Let

$$\hat{\Phi}(\mathbf{X}) = (\hat{D}_1 X_{2-q_1} - X_{1+q_1}\hat{C}_1,\ldots,\hat{D}_p X_{p+1-q_p} - X_{p+q_p}\hat{C}_p) \tag{14}$$

and introduce the quadratic transformation

$$\Psi(\mathbf{X}) := (X^H_{1+q_1}\hat{F}_1 X_{2-q_1},\ldots,X^H_{p+q_p}\hat{F}_p X_{p+1-q_p}), \tag{15}$$

as well as the tuple $\mathbf{E} = (E_1,\ldots,E_p)$. Then (13) can be rewritten as

$$\hat{\Phi}(\mathbf{X}) - \mathbf{E} + \Psi(\mathbf{X}) = 0. \tag{16}$$

If a solution $\mathbf{X}$ to (16) exists, then a straightforward calculation yields

$$Y_{k+q_k}^H \hat{T}_k Y_{k+1-q_k} = \tag{17}$$
$$\begin{bmatrix} H_{1,k+q_k}^{-1}(\hat{C}_k - \hat{F}_k X_{k+1-q_k}) H_{1,k+1-q_k} & * \\ 0 & H_{2,k+q_k}(\hat{D}_k + X_{k+q_k}^H \hat{F}_k) H_{2,k+1-q_k}^{-1} \end{bmatrix}.$$

To evaluate upper bounds of $\|X_1\|, \ldots, \|X_p\|$, we introduce a norm on matrix tuples $\mathbf{X} = (X_1, \ldots, X_p)$ via

$$|||\mathbf{X}||| := \max_{k \in \{1,\ldots,p\}} \|X_k\|.$$

For $\hat{\Phi}(X)$ as in (14) we set

$$\hat{\delta} := \min_{|||\mathbf{X}|||=1} |||\hat{\Phi}(\mathbf{X})||| \tag{18}$$

and similarly for $\Phi(\mathbf{X})$ as in (7)

$$\delta := \min_{|||\mathbf{X}|||=1} |||\Phi(\mathbf{X})|||. \tag{19}$$

The quantities δ and $\hat{\delta}$ are generalizations of the sep operator for matrices and matrix pencils, see [15, 24]. Since the quantities $\mathbf{C}$, $\mathbf{D}$ and the perturbed quantities $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$ are related via (7), (10), (11), and (14), we have the following inequalities

$$\delta - |||\Delta\mathbf{C}||| - |||\Delta\mathbf{D}||| \leq \hat{\delta} \leq \delta + |||\Delta\mathbf{C}||| + |||\Delta\mathbf{D}|||. \tag{20}$$

For $\Psi(\mathbf{X})$ as in (15), using the tuple $\hat{\mathbf{F}} = (\hat{F}_1, \ldots, \hat{F}_p)$, we obtain

$$|||\Psi(\mathbf{X})||| \leq |||\hat{\mathbf{F}}||| \, |||\mathbf{X}|||^2 \tag{21}$$

and

$$|||\Psi(\mathbf{X}) - \Psi(\mathbf{Y})||| \leq 2|||\hat{\mathbf{F}}||| \max\{|||\mathbf{X}|||, |||\mathbf{Y}|||\} |||\mathbf{X} - \mathbf{Y}|||. \tag{22}$$

We then have the following perturbation result.

Theorem 3 *Let* $\mathbf{T}$ *be as in (6),* $\hat{\mathbf{T}} = \mathbf{T} + \Delta\mathbf{T}$ *as in (11),* $\Delta\mathbf{T}$ *as in (10),* Φ *as in (7),* $\hat{\Phi}$ *as in (14), and* Ψ *be as in (15). If* $\hat{\delta} > 0$ *is as in (18) and if*

$$\frac{|||\mathbf{E}||| \, |||\hat{\mathbf{F}}|||}{\hat{\delta}^2} < \frac{1}{4}, \tag{23}$$

then there exists a unique solution $\mathbf{X} = (X_1, \ldots, X_p)$ *of (16) satisfying*

$$|||\mathbf{X}||| \leq \frac{2|||\mathbf{E}|||}{\hat{\delta} + \sqrt{\hat{\delta}^2 - 4|||\hat{\mathbf{F}}||| \, |||\mathbf{E}|||}} < 2\frac{|||\mathbf{E}|||}{\hat{\delta}}. \tag{24}$$

Proof. Since the transformation Ψ satisfies (21) and (22), and since $\hat{\delta} > 0$ the result follows from Theorem V.2.11 in [24, p.242] together with (23), applied to the quadratic equation (16). □

Using this result we get the following perturbation result for generalized deflating subspaces of $\mathbf{A}$.

Theorem 4 *Let* $\mathbf{A} = (A_1, \dots, A_p)$ *be a regular tuple of the form (6) with sign tuple* $s = (s_1, \dots, s_p)$. *Let* $Q_k = [U_k, V_k]$, *for* $k = 1, \dots, p$, *and let* $\mathbf{U} = (U_1, \dots, U_p)$ *be an orthonormal basis of the generalized deflating subspace corresponding to* $\Lambda(\mathbf{C})$. *Let* $\hat{\mathbf{A}} = (A_1 + \Delta A_1, \dots, A_p + \Delta A_p)$ *be the perturbed matrix tuple and let* $\Delta\mathbf{T} = (\Delta T_1, \dots, \Delta T_p)$ *with* $\Delta T_k = Q_{k+q_k}^H \Delta A_k Q_{k+1-q_k}$ *be partitioned as in (10). If* $\hat{\delta} > 0$ *satisfies (23), then* $\hat{\mathbf{A}}$ *has a generalized deflating subspace with orthonormal basis*

$$\hat{\mathbf{U}} := (\hat{U}_1, \dots, \hat{U}_p) = (Q_1 \begin{bmatrix} I_m \\ -X_1 \end{bmatrix} H_{11}, \dots, Q_p \begin{bmatrix} I_m \\ -X_p \end{bmatrix} H_{1p}) \tag{25}$$

corresponding to the eigenvalues of

$$\tilde{\mathbf{C}} = (H_{1,1+q_1}^{-1}(\hat{C}_1 - \hat{F}_1 X_{2-q_1}) H_{1,2-q_1}, \dots, H_{1,p+q_p}^{-1}(\hat{C}_p - \hat{F}_p X_{p+1-q_p}) H_{1,p+1-q_p}), \tag{26}$$

where $H_{1k} = (I_m + X_k^H X_k)^{-\frac{1}{2}}$ *for* $k = 1, \dots, p$.

Furthermore, for $k = 1, \dots, p$ *and* $\hat{\delta}$ *as in (18), the maximal principal angle between* $\operatorname{range} U_k$ *and* $\operatorname{range} \hat{U}_k$ *is less than* $\arctan(2\frac{\|\mathbf{E}\|}{\hat{\delta}})$.

Proof. Relations (25) and (26) follow from the relationship between $\mathbf{A}$, $\mathbf{T}$ and the perturbed quantities $\hat{\mathbf{A}}$, $\hat{\mathbf{T}}$, respectively, Theorem 3 and formula (17).

Following ([24, Corollary I.5.4]) the principal angle between $\operatorname{range} U_k$ and $\operatorname{range} \hat{U}_k$ is given by

$$\arcsin \|V_k^H \hat{U}_k\| = \arcsin \|X_k H_{1k}\| = \arcsin \frac{\|X_k\|}{\sqrt{1 + \|X_k\|^2}}.$$

Using the monotonicity of the function $\arcsin \frac{\xi}{\sqrt{1+\xi^2}}$ and the fact that $\|X_k\| \le \|\mathbf{X}\|$, the last statement follows. □

Using (20), the conditions $\hat{\delta} > 0$ and (23) in Theorem 3 can be replaced by

$$\rho := \delta - \|\Delta\mathbf{C}\| - \|\Delta\mathbf{D}\| > 0 \tag{27}$$

and

$$\frac{\|\mathbf{E}\|(\|\mathbf{F}\| + \|\Delta\mathbf{F}\|)}{\rho^2} < \frac{1}{4}, \tag{28}$$

respectively. In this case we obtain

$$\|\mathbf{X}\| \le \frac{2\|\mathbf{E}\|}{\rho + \sqrt{\rho^2 - 4\|\mathbf{E}\|(\|\mathbf{F}\| + \|\Delta\mathbf{F}\|)}} < 2\frac{\|\mathbf{E}\|}{\rho}.$$

Remark 1 By definition, $\delta > 0$ is a necessary and sufficient condition for the nonsingularity of Φ. Since $\hat{\delta} > 0$ we obtain that $\hat{\Phi}$ is nonsingular and $\Lambda(\hat{\mathbf{C}}) \cap \Lambda(\hat{\mathbf{D}}) = \emptyset$. Similarly, using (20), condition (27) implies that both Φ and $\hat{\Phi}$ are nonsingular, $\Lambda(\mathbf{C}) \cap \Lambda(\mathbf{D}) = \emptyset$ and $\Lambda(\hat{\mathbf{C}}) \cap \Lambda(\hat{\mathbf{D}}) = \emptyset$.

Remark 2 The conditions $\hat{\delta} > 0$ and (23) imply that $\Lambda(\tilde{\mathbf{C}}) \cap \Lambda(\tilde{\mathbf{D}}) = \emptyset$, where $\tilde{\mathbf{C}}$ is in (26) and

$$\tilde{\mathbf{D}} := (H_{2,1+q_1}(\hat{D}_1 + X_{1+q_1}^H \hat{F}_1) H_{2,2-q_1}^{-1}, \dots, H_{2,p+q_p}(\hat{D}_p + X_{p+q_p}^H \hat{F}_p) H_{2,p+1-q_p}^{-1}),$$

where $H_{2k} = (I_{n-m} + X_k X_k^H)^{-\frac{1}{2}}$, for $k = 1, \ldots, p$. To show this, by Lemma 1 and Remark 1 we only need to show that for the linear transformation $\Phi_{\tilde{\mathbf{C}},\tilde{\mathbf{D}}}$ corresponding to $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{D}}$ we have

$$\min_{\|\mathbf{Z}\|=1} \|\Phi_{\tilde{\mathbf{C}},\tilde{\mathbf{D}}}(\mathbf{Z})\| > 0.$$

Using inequalities similar to (18), (23) and (24) it is straightforward to show that

$$\min_{\|\mathbf{Z}\|=1} \Phi_{\tilde{\mathbf{C}},\tilde{\mathbf{D}}}(\mathbf{Z}) \geq \frac{1}{1+\|\mathbf{X}\|^2}\left(\hat{\delta} - \frac{4\|\mathbf{E}\|\|\hat{\mathbf{F}}\|}{\hat{\delta}}\right) > 0.$$

Similar bounds are also obtained if the conditions $\hat{\delta} > 0$ and (23) are replaced by $\rho > 0$ and (28), respectively.

Remark 3 The quantity δ can be considered as the reciprocal of the condition number for the generalized deflating subspace. If the norm is the spectral norm, then it is difficult to evaluate δ. However, if we use the Frobenius norm, and the induced norm $\|\mathbf{X}\|_F := \|[X_1, \ldots, X_p]\|_F = \sqrt{\sum_{k=1}^{p} \|X_k\|_F^2}$, we can determine

$$\delta_F := \min_{\|\mathbf{X}\|_F=1} \|\Phi(\mathbf{X})\|_F = \sigma_{\min}(Z),$$

where the matrix Z is defined in (8).

The results of this section show that the classical perturbation results for deflating subspaces as in [24] can be extended to generalized deflating subspaces. In the next subsection we derive perturbation results for simple eigenvalues in a similar way.

2.2 Eigenvalue perturbations

In this subsection we study the first order perturbation analysis of simple eigenvalues and the associated eigenvectors of formal matrix products for sufficiently small perturbations $\Delta\mathbf{A} = (\Delta A_1, \ldots, \Delta A_p)$.

Theorem 5 *Consider a regular tuple* $\mathbf{A}$ *with sign tuple* s *and let* λ *be a simple eigenvalue of* $\mathbf{A}$. *Let* $\mathbf{A}$ *be transformed via*

$$Q_{k+q_k}^H A_k Q_{k+1-q_k} = \begin{bmatrix} \alpha_k & F_k \\ 0 & D_k \end{bmatrix} =: T_k, \tag{29}$$

for $k = 1, \ldots, p$. *Let* $\lambda = \alpha_1^{s_1} \cdots \alpha_p^{s_p}$ *and let* $\mathbf{u} = (Q_1 e_1, \ldots, Q_p e_1)$ *be the unit norm right eigenvector associated with* λ. *Consider a perturbed tuple* $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}$ *and set*

$$Q_{k+q_k}^H \Delta A_k Q_{k+1-q_k} = \begin{bmatrix} \Delta\alpha_k & \Delta F_k \\ E_k & \Delta D_k \end{bmatrix}.$$

If $\|\Delta\mathbf{A}\|$ *is sufficiently small, then there exists a unit norm right eigenvector* $\hat{\mathbf{u}}$ *of* $\hat{\mathbf{A}}$ *with* $\hat{A}_k \hat{u}_{k+1-q_k} = \hat{\alpha}_k \hat{u}_{k+q_k}$, *such that for* $k = 1, \ldots, p$,

$$\hat{\alpha}_k - \alpha_k = \Delta\alpha_k - F_k x_{k+1-q_k} + O(\|\Delta\mathbf{A}\|^2), \tag{30}$$

$$\hat{u}_k - u_k = Q_k \begin{bmatrix} 0 \\ -x_k \end{bmatrix} + O(\|\Delta\mathbf{A}\|^2). \tag{31}$$

Here the vectors $x_1, \ldots, x_p$ solve the equations

$$D_k x_{k+1-q_k} - \alpha_k x_{k+q_k} = E_k, \quad k = 1, \ldots, n.$$

Moreover, let δ be defined as in (19), then for $k = 1, \ldots, p$

$$|\hat{\alpha}_k - \alpha_k| \leq |\!|\!|\Delta \mathbf{A}|\!|\!|(1 + \frac{|\!|\!|\mathbf{F}|\!|\!|}{\delta}) + O(|\!|\!|\Delta \mathbf{A}|\!|\!|^2), \tag{32}$$

and

$$|\!|\!|\mathbf{u} - \hat{\mathbf{u}}|\!|\!| \leq \frac{|\!|\!|\Delta \mathbf{A}|\!|\!|}{\delta} + O(|\!|\!|\Delta \mathbf{A}|\!|\!|^2). \tag{33}$$

Proof. Since the eigenvector is the simplest case of a generalized deflating subspace, equations (30) and (31) follow directly from Theorem 4 and (17). Since $|\!|\!|\Delta \mathbf{A}|\!|\!|$ is sufficiently small, separating the first order perturbations, equation (16) can be written as

$$\Phi(\mathbf{X}) = \mathbf{E} = (E_1, \ldots, E_p) + O(|\!|\!|\Delta \mathbf{A}|\!|\!|^2).$$

Hence

$$|\!|\!|\mathbf{X}|\!|\!| \leq \frac{|\!|\!|\mathbf{E}|\!|\!|}{\delta} + O(|\!|\!|\Delta \mathbf{A}|\!|\!|^2) \leq \frac{|\!|\!|\Delta \mathbf{A}|\!|\!|}{\delta} + O(|\!|\!|\Delta \mathbf{A}|\!|\!|^2)$$

and the bounds (32), (33) follow. □

Note that we have given the perturbations in the components α_k rather than in λ itself. But since the factors $\alpha_1, \ldots, \alpha_p$ are uniquely determined up to a unit modular factor in each α_k, (30) immediately gives a first order perturbation bound for the eigenvalue λ, too. However, we will also give a different expression by employing the left eigenvector. For this we use (29) and obtain a linear system for the vectors $z_1, \ldots, z_p$ as

$$\alpha_k z_{k+1-q_k}^H - z_{k+q_k}^H D_k = F_k, \quad k = 1, \ldots, p. \tag{34}$$

If λ is a simple eigenvalue, then as for Lemma 1 we can show that the linear transformation corresponding to the left side of (34) is nonsingular. Hence (34) has a unique solution $z_1, \ldots, z_p$. Now set

$$\tilde{w}_k = \begin{bmatrix} 1 \\ -z_k \end{bmatrix}, \quad w_k = \frac{\tilde{w}_k}{\|\tilde{w}_k\|}, \quad \mathbf{w} := (w_1, \ldots, w_p).$$

Then the vectors $w_1, \ldots, w_p$ have unit norm and satisfy

$$w_{k+q_k}^H T_k = \beta_k w_{k+1-q_k}^H, \quad \beta_k = \alpha_k \frac{\|\tilde{w}_{k+1-q_k}\|}{\|\tilde{w}_{k+q_k}\|}, \quad k = 1, \ldots, p.$$

Hence $\mathbf{w}$ can be viewed as a unit norm *left eigenvector* of $\mathbf{T}$ corresponding to λ. Note that the related unit norm right eigenvector of $\mathbf{T}$ now is $(e_1, \ldots, e_1)$ and we have

$$w_k^H e_1 = \frac{1}{\|\tilde{w}_k\|} > 0, \quad k = 1, \ldots, p. \tag{35}$$

Obviously, $(Q_1 w_1, \ldots, Q_p w_p)$ is the unit norm left eigenvector of $\mathbf{A}$. Similar to Lemma 2, for λ simple, we can show that the unit norm left eigenvector $\mathbf{v} = (v_1, \ldots, v_p)$ of $\mathbf{A}$ corresponding to λ is unique and satisfies

$$v_{k+q_k}^H A_k = \beta_k v_{k+1-q_k}^H, \quad k = 1, \ldots, p, \tag{36}$$

where $\lambda = \beta_1^{s_1} \cdots \beta_p^{s_p}$. Let $\mathbf{u} = (u_1, \ldots, u_p)$ be a corresponding unit norm right eigenvector, i.e.,

$$A_k u_{k+1-q_k} = \alpha_k u_{k+q_k}, \quad k = 1, \ldots, p, \quad \lambda = \alpha_1^{s_1} \cdots \alpha_p^{s_p}. \tag{37}$$

Multiplying u_{k+1-q_k} from the right-hand side of (36) we obtain

$$\alpha_k v_{k+q_k}^H u_{k+q_k} = \beta_k v_{k+1-q_k}^H u_{k+1-q_k}. \tag{38}$$

Due to the uniqueness of the eigenvector, it follows from (35) that

$$\kappa_k := v_k^H u_k \neq 0, \quad k = 1, \ldots, p. \tag{39}$$

Note that if all the matrices $B_1, \ldots, B_p$ in (3) exist, one can verify that $|\kappa_1|, \ldots, |\kappa_p|$ are just the reciprocal condition numbers corresponding to the same eigenvalue λ of $B_1, \ldots, B_p$ respectively, and hence the classical condition number for a simple eigenvalue of one matrix, see [26], is reproduced.

Using these relations we now derive the first order perturbation theory for the eigenvalues of a perturbed formal product. For this we need to separate the positive and negative signs in the sign tuple $s = (s_1, \ldots, s_p)$ via

$$I_+ = \{k \mid s_k = 1\}, \quad I_- = \{k \mid s_k = -1\}.$$

Theorem 6 *Consider a regular tuple $\mathbf{A}$ with sign tuple s of the form (29). Let $\lambda \in \Lambda(\mathbf{A})$ be a simple eigenvalue and let*

$$\mathbf{u} = (u_1, \ldots, u_p), \quad \mathbf{v} = (v_1, \ldots, v_p)$$

be the corresponding unit norm right and left eigenvectors satisfying (36) and (37), respectively. Let $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}$ with $\|\Delta\mathbf{A}\|$ sufficiently small. Then the perturbed tuple $\hat{\mathbf{A}}$ has unit norm eigenvectors $\hat{\mathbf{u}} = (\hat{u}_1, \ldots, \hat{u}_p)$ satisfying

$$\hat{A}_k \hat{u}_{k+1-q_k} = \hat{\alpha}_k \hat{u}_{k+q_k}, \tag{40}$$

for $k = 1, \ldots, p$, such that with κ_k defined in (39) the perturbations satisfy

$$\prod_{k \in I_+} \hat{\alpha}_k \prod_{k \in I_-} \alpha_k - \prod_{k \in I_-} \hat{\alpha}_k \prod_{k \in I_+} \alpha_k = \sum_{j=1}^{p} (-1)^{q_j} \left(\prod_{k \neq j} \alpha_k \right) \frac{v_{j+q_j}^H \Delta A_j u_{j+1-q_j}}{\kappa_{j+q_j}} + O(\|\Delta\mathbf{A}\|^2). \tag{41}$$

Proof. Expansions (30) and (31) in Theorem 5 imply that

$$\begin{aligned} \hat{\alpha}_k &= \alpha_k + \Delta\alpha_k + O(\|\Delta\mathbf{A}\|^2), \quad |\Delta\alpha_k| = O(\|\Delta\mathbf{A}\|), \\ \hat{u}_k &= u_k + \Delta u_k + O(\|\Delta\mathbf{A}\|^2), \quad \|\Delta u_k\| = O(\|\Delta\mathbf{A}\|), \end{aligned}$$

for $k = 1, \ldots, p$. Using these expansions in (40) and applying (37), it follows that the first order terms satisfy

$$A_k \Delta u_{k+1-q_k} + \Delta A_k u_{k+1-q_k} = \alpha_k \Delta u_{k+q_k} + \Delta\alpha_k u_{k+q_k}.$$

Multiplying by $v_{k+q_k}^H$ from the left and using (36) and (39) we then get

$$\Delta\alpha_k = \frac{1}{\kappa_{k+q_k}} (\beta_k v_{k+1-q_k}^H \Delta u_{k+1-q_k} - \alpha_k v_{k+q_k}^H \Delta u_{k+q_k} + v_{k+q_k}^H \Delta A_k u_{k+1-q_k}) + O(\|\Delta\mathbf{A}\|^2).$$

Using the relation

$$\frac{\beta_k}{\kappa_{k+q_k}} = \frac{\alpha_k}{\kappa_{k+1-q_k}},$$

which is from (38), we have

$$\hat{\alpha}_k - \alpha_k = \frac{v_{k+q_k}^H \Delta A_k u_{k+1-q_k}}{\kappa_{k+q_k}} + \alpha_k \left(\frac{v_{k+1-q_k}^H \Delta u_{k+1-q_k}}{\kappa_{k+1-q_k}} - \frac{v_{k+q_k}^H \Delta u_{k+q_k}}{\kappa_{k+q_k}} \right) + O(\|\Delta \mathbf{A}\|^2), \quad (42)$$

for $k = 1, \ldots, p$. Note that $q_k = 0$ if $s_k = 1$ and $q_k = 1$ if $s_k = -1$. Expansion (42) then implies that

$$\begin{aligned} \prod_{k \in I_+} \hat{\alpha}_k &= \left(\prod_{k \in I_+} \alpha_k \right) \left(1 + \sum_{k \in I_+} \left(\frac{v_{k+1}^H \Delta u_{k+1}}{\kappa_{k+1}} - \frac{v_k^H \Delta u_k}{\kappa_k} \right) \right) \\ &\quad + \sum_{j \in I_+} \left(\prod_{k \in I_+, k \neq j} \alpha_k \right) \frac{v_j^H \Delta A_j u_{j+1}}{\kappa_j} + O(\|\Delta \mathbf{A}\|^2) \end{aligned}$$

and similarly

$$\begin{aligned} \prod_{k \in I_-} \hat{\alpha}_k &= \left(\prod_{k \in I_-} \alpha_k \right) \left(1 + \sum_{k \in I_-} \left(\frac{v_k^H \Delta u_k}{\kappa_k} - \frac{v_{k+1}^H \Delta u_{k+1}}{\kappa_{k+1}} \right) \right) \\ &\quad + \sum_{j \in I_-} \left(\prod_{k \in I_-, k \neq j} \alpha_k \right) \frac{v_{j+1}^H \Delta A_j u_j}{\kappa_{j+1}} + O(\|\Delta \mathbf{A}\|^2). \end{aligned}$$

Using the periodicity, i.e., that

$$\frac{v_{p+1}^H \Delta u_{p+1}}{\kappa_{p+1}} = \frac{v_1^H \Delta u_1}{\kappa_1},$$

the identity

$$\sum_{k=1}^{p} \left(\frac{v_k^H \Delta u_k}{\kappa_k} - \frac{v_{k+1}^H \Delta u_{k+1}}{\kappa_{k+1}} \right) = 0$$

implies that

$$\sum_{k \in I_-} \left(\frac{v_k^H \Delta u_k}{\kappa_k} - \frac{v_{k+1}^H \Delta u_{k+1}}{\kappa_{k+1}} \right) = \sum_{k \in I_+} \left(\frac{v_{k+1}^H \Delta u_{k+1}}{\kappa_{k+1}} - \frac{v_k^H \Delta u_k}{\kappa_k} \right).$$

Hence

$$\begin{aligned} &\prod_{k \in I_+} \hat{\alpha}_k \prod_{k \in I_-} \alpha_k - \prod_{k \in I_-} \hat{\alpha}_k \prod_{k \in I_+} \alpha_k \\ &= \sum_{j \in I_+} \left(\prod_{k \neq j} \alpha_k \right) \frac{v_j^H \Delta A_j u_{j+1}}{\kappa_j} - \sum_{j \in I_-} \left(\prod_{k \neq j} \alpha_k \right) \frac{v_{j+1}^H \Delta A_j u_j}{\kappa_{j+1}} + O(\|\Delta \mathbf{A}\|^2) \\ &= \sum_{j=1}^{p} (-1)^{q_j} \left(\prod_{k \neq j} \alpha_k \right) \frac{v_{j+q_j}^H \Delta A_j u_{j+1-q_j}}{\kappa_{j+q_j}} + O(\|\Delta \mathbf{A}\|^2), \end{aligned}$$

which is (41). □

Expansion (41) gives first order perturbations only for the α_k, but the first order perturbations for λ are easily derived as a Corollary.

Corollary 7 *Consider a regular tuple* $\mathbf{A}$ *with sign tuple* s*. Let* λ *be a simple eigenvalue of the formal product and let* $\alpha_1, \ldots, \alpha_p$ *associated with* λ *satisfy (37). If* $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}$ *with* $\|\Delta\mathbf{A}\|$ *sufficiently small, then* $\hat{\mathbf{A}}$ *has an eigenvalue* $\hat{\lambda}$ *that satisfies the following first order perturbation results.*

a) *If* λ *is finite and nonzero then*

$$\frac{\hat{\lambda} - \lambda}{\lambda} = \sum_{k=1}^{p} (-1)^{q_k} \frac{v_{k+q_k}^H \Delta A_k u_{k+1-q_k}}{\alpha_k \kappa_{k+q_k}} + O(\|\Delta\mathbf{A}\|^2). \tag{43}$$

b) *If* $\lambda = 0$ *and* k_0 *is an index such that* $s_{k_0} = 1$ *and* $\alpha_{k_0} = 0$*, then*

$$\hat{\lambda} = \frac{\prod_{k \in I_+, k \neq k_0} \alpha_k}{\prod_{k \in I_-} \alpha_k} \frac{v_{k_0}^H \Delta A_{k_0} u_{k_0+1}}{\kappa_{k_0}} + O(\|\Delta\mathbf{A}\|^2). \tag{44}$$

c) *If* $\lambda = \infty$ *and* k_0 *is an index such that* $s_{k_0} = -1$ *and* $\alpha_{k_0} = 0$*, then*

$$\frac{1}{\hat{\lambda}} = -\frac{\prod_{k \in I_-, k \neq k_0} \alpha_k}{\prod_{k \in I_+} \alpha_k} \frac{v_{k_0+1}^H \Delta A_{k_0} u_{k_0}}{\kappa_{k_0+1}} + O(\|\Delta\mathbf{A}\|^2). \tag{45}$$

Proof. If λ is finite and nonzero then all α_k are nonzero. If $\|\Delta\mathbf{A}\|$ is sufficiently small, then from (32) we also have that $\prod_{k \in I_-} \hat{\alpha}_k \neq 0$. Multiplying with

$$\frac{1}{\lambda \left(\prod_{k \in I_-} \alpha_k\right) \left(\prod_{k \in I_-} \hat{\alpha}_k\right)}$$

on both sides of (41) and using that

$$\lambda = \frac{\prod_{k \in I_+} \alpha_k}{\prod_{k \in I_-} \alpha_k}, \qquad \prod_{k \in I_-} \hat{\alpha}_k = \prod_{k \in I_-} \alpha_k + O(\|\Delta\mathbf{A}\|),$$

we obtain (43).

If $\lambda = 0$ then, since $\mathbf{A}$ is regular, there exists at least one $k_0 \in I_+$ such that $\alpha_{k_0} = 0$ and $\prod_{k \in I_-} \alpha_k \neq 0$. Hence the right side of (41) reduces to

$$\left(\prod_{k \neq k_0} \alpha_k\right) \frac{v_{k_0}^H \Delta A_{k_0} u_{k_0+1}}{\kappa_{k_0}} + O(\|\Delta\mathbf{A}\|^2).$$

Similarly, by multiplying with

$$\frac{1}{\left(\prod_{k \in I_-} \alpha_k\right) \left(\prod_{k \in I_-} \hat{\alpha}_k\right)}$$

on both sides of (41) we obtain (44).

The expansion (45) in the case $\lambda = \infty$ is derived similarly to the case $\lambda = 0$. $\square$

We see that the perturbations for an eigenvalue λ and its components α_k are of slightly different nature. For the component α_k from (42) the perturbation has two parts. One arises directly from ΔA_k in the term

$$\frac{v_{k+q_k}^H \Delta A_k u_{k+1-q_k}}{\kappa_{k+q_k}}.$$

The other part arises from the perturbation of the eigenvector in the term

$$\alpha_k \left(\frac{v_{k+1-q_k}^H \Delta u_{k+1-q_k}}{\kappa_{k+1-q_k}} - \frac{v_{k+q_k}^H \Delta u_{k+q_k}}{\kappa_{k+q_k}} \right).$$

For the eigenvalue λ, however, only the first term occurs. But, nevertheless, we see from (30) and (41) that the perturbations in λ and α_k are of the same order.

Remark 4 Corollary 7 implies that for the eigenvalue 0 with at least two indices k_1, k_2 such that $\alpha_{k_1} = \alpha_{k_2} = 0$ with $s_{k_1} = s_{k_2} = 1$, the corresponding perturbed eigenvalue is of second order. The same holds for the eigenvalue infinity if there exists $\alpha_{k_1} = \alpha_{k_2} = 0$ with $s_{k_1} = s_{k_2} = -1$.

In this subsection we have shown that the classical perturbation results for simple eigenvalues and associated eigenvectors can be directly extended to formal matrix products. The perturbations for the factors of an eigenvalue are slightly different from those for the complete factor as was to be expected already from the perturbation theory of matrix pencils, see [24]. The situation changes drastically for the case of multiple eigenvalues that we discuss in the next subsection.

2.3 Perturbations of multiple eigenvalues

Even for the case of one matrix the perturbation theory for multiple eigenvalues is complicated. If the matrix is diagonalizable, then the perturbations for the eigenvalue still are similar to those for simple eigenvalues [26]. However, for the eigenvectors usually there are no similar results. For completeness we will present the perturbation result for multiple eigenvalues with a slightly different proof than in [24]. This proof will then also be used for the formal matrix product case.

Theorem 8 *Let $A \in \mathbb{C}^{n \times n}$ be diagonalizable, let λ be an eigenvalue of A of algebraic multiplicity m and let U, V form orthonormal bases of the corresponding right and left eigenvector spaces. Consider a perturbation $\hat{A} = A + \Delta A$ with $\|\Delta A\|$ sufficiently small. Then $\hat{A}$ has m associated eigenvalues and for each such eigenvalue $\hat{\lambda}$, there exists a unit norm eigenvector $x \in \mathbb{C}^m$. With an arbitrary nonzero vector $y \in \mathbb{C}^m$ such that $y^H V^H U x \neq 0$ we have*

$$\hat{\lambda} - \lambda = \frac{y^H V^H \Delta A U x}{y^H V^H U x} + O(\|\Delta A\|^2). \tag{46}$$

Moreover,

$$\begin{aligned} |\hat{\lambda} - \lambda| &= \min_y \left| \frac{y^H V^H \Delta A U x}{y^H V^H U x} \right| + O(\|\Delta A\|^2) \\ &\leq \frac{\|\Delta A\|}{\|V^H U x\|} + O(\|\Delta A\|^2) \leq \frac{\|\Delta A\|}{\sigma_{\min}(V^H U)} + O(\|\Delta A\|^2). \end{aligned}$$

Proof. By assumption there exists a unitary matrix Q with $Q = [U, \tilde{U}]$ such that

$$AQ = Q \begin{bmatrix} \lambda I_m & F \\ 0 & D \end{bmatrix}.$$

Partition

$$Q^H \hat{A} Q = \begin{bmatrix} \lambda I_m + \Delta C & F + \Delta F \\ E & D + \Delta D \end{bmatrix}.$$

Since $\|\Delta A\|$ is sufficiently small, there exists a matrix X solving

$$(D + \Delta D)X - X(\lambda I_m + \Delta C) - E + X(F + \Delta F)X = 0,$$

and X is of order $\|\Delta A\|$. Then

$$\hat{A} Q \begin{bmatrix} I \\ -X \end{bmatrix} = Q \begin{bmatrix} I \\ -X \end{bmatrix} (\lambda I + \Delta C - (F + \Delta F)X). \tag{47}$$

Let $\Delta\lambda$ be an eigenvalue of $\Delta C - (F + \Delta F)X$ with corresponding unit norm eigenvector x. Clearly $\Delta\lambda$ is of order $\|\Delta A\|$ and $\lambda + \Delta\lambda$ is an eigenvalue of $\hat{A}$. Pre- and postmultiplying $y^H V^H$, x in (47) and using the formulas for V and Q, if $y^H V^H U x \neq 0$, then we get

$$y^H V^H \Delta A U x = \Delta\lambda\, y^H V^H U x + O(\|\Delta A\|^2)$$

and we obtain (46). Setting $y = \frac{1}{\|V^H U x\|} V^H U x$ we have the first upper bound. The second bound follows from $\|V^H U x\| \geq \sigma_{\min}(V^H U)$. $\square$

As in the classical case of matrices and pencils, the reciprocal of the condition number of a multiple eigenvalue λ is given by $\sigma_{\min}(V^H U)$.

Unlike for the case of simple eigenvalues, the eigenvectors in (46) depend on the perturbations. Neither the eigenvalues nor the eigenvectors are analytic functions in the elements of ΔA in the neighborhood of the origin. For example, let $A = I_2$ and $\Delta A = \left[\begin{smallmatrix} \epsilon & \epsilon \\ \epsilon^2 & \epsilon \end{smallmatrix}\right]$. Then $\hat{A}$ has two eigenvalues $1 + \epsilon \pm |\epsilon|^{\frac{3}{2}}$. It may also happen that the perturbed matrix is not diagonalizable, as we see from the example $A = I_2$ and $\Delta A = \left[\begin{smallmatrix} \epsilon & \epsilon \\ 0 & \epsilon \end{smallmatrix}\right]$.

For a matrix tuple $\mathbf{A}$ with sign tuple s, let λ be an eigenvalue of $\mathbf{A}$ with algebraic multiplicity m. If there exists a matrix tuple $\mathbf{W} = (W_1, \ldots, W_p)$ with $W_k \in \mathbb{C}^{n \times m}$ of full column rank such that

$$A_k W_{k+1-q_k} = W_{k+q_k} \Gamma_k, \quad \Gamma_k = \mathrm{diag}(\gamma_{1k}, \ldots, \gamma_{mk}),$$

for $k = 1, \ldots, p$ and $\lambda = \prod_{k=1}^p \gamma_{1;k}^{s_k} = \ldots = \prod_{k=1}^p \gamma_{m;k}^{s_k}$, ($\prod_{k=1}^p \gamma_{1;k}^{-s_k} = \ldots = \prod_{k=1}^p \gamma_{m;k}^{-s_k} = 0$ for infinite eigenvalues), then we say that λ has *a complete set of right eigenvectors.*

Let U_k be an orthonormal basis of range W_k and let $Q_k = [U_k, \tilde{U}_k]$ be unitary. As before we set

$$Q_{k+q_k}^H A_k Q_{k+1-q_k} = \begin{bmatrix} C_k & F_k \\ 0 & D_k \end{bmatrix}, \tag{48}$$

and

$$A_k U_{k+1-q_k} = U_{k+q_k} C_k, \tag{49}$$

with $\Lambda(\mathbf{C}) = \{\lambda\}$. Moreover, if λ is finite, then

$$C_k^{s_k} \cdots C_p^{s_p} C_1^{s_1} \cdots C_{k-1}^{s_{k-1}} = \lambda I_m, \tag{50}$$

and if λ is infinite, then

$$C_k^{-s_k} \cdots C_p^{-s_p} C_1^{-s_1} \cdots C_{k-1}^{-s_{k-1}} = 0, \tag{51}$$

for all $k = 1, \ldots, p$. If λ is nonzero finite, then all C_k are nonsingular and we can verify that (50) holds for all $k = 1, \ldots, p$ if and only if it holds for one k. Moreover, (50) is also a sufficient condition for $\mathbf{A}$ to have a complete set of eigenvectors associated with an eigenvalue λ. To verify this one can simply take $W_1 = U_1$ and $W_k = U_k \prod_{j=k}^{p} C_j^{s_j}$ for $k = 2, \ldots, p$. If λ is zero or infinite, however, then we do not know of such a simple connection. We conjecture that if equations (50) or (51) hold for all $k = 1, \ldots, p$ then also complete sets of eigenvectors exist for the eigenvalues zero and infinity.

We will now analyze perturbations in equations (49) and (50), (51). Let $\mathbf{V}$ be an orthonormal basis of the left eigenvector subspace, i.e.,

$$V_{k+q_k}^H A_k = \tilde{C}_k V_{k+1-q_k}^H. \tag{52}$$

Then as shown in Subsection 2.1, $H_k := V_k^H U_k$ is nonsingular, and from (52) and (49), we obtain

$$\tilde{C}_k = H_{k+q_k} C_k H_{k+1-q_k}^{-1}, \quad k = 1, \ldots, p. \tag{53}$$

Let $\mathbf{A} + \Delta\mathbf{A}$ be the perturbed matrix tuple with $\|\Delta\mathbf{A}\|$ sufficiently small. Then as in Subsection 2.1 there exists $\mathbf{X}$ with $\|\mathbf{X}\| = O(\|\Delta\mathbf{A}\|)$ such that for $\hat{U}_k := Q_k \begin{bmatrix} I \\ -X_k \end{bmatrix}$ we have

$$(A_k + \Delta A_k)\hat{U}_{k+1-q_k} = \hat{U}_{k+q_k}(C_k + \Delta C_k), \tag{54}$$

where

$$\Delta C_k = U_{k+q_k}^H \Delta A_k U_{k+1-q_k} + (F_k + U_{k+q_k}^H \Delta A_k \tilde{U}_{k+1-q_k}) X_{k+1-q_k},$$

for $k = 1, \ldots, p$. As $\|\Delta\mathbf{C}\| = O(\|\Delta\mathbf{A}\|)$ and $\|\Delta\mathbf{A}\|$ is assumed to be sufficiently small, the eigenvalues of $\mathbf{C}+\Delta\mathbf{C}$ are just the m eigenvalues of $\mathbf{A}+\Delta\mathbf{A}$ nearest to λ. Let $\mathbf{x} = (x_1, \ldots, x_p)$ be the unit norm right eigenvector of an eigenvalue of $\mathbf{C} + \Delta\mathbf{C}$, i.e.,

$$(C_k + \Delta C_k)x_{k+1-q_k} = \hat{\alpha}_k x_{k+q_k}, \quad k = 1, \ldots, p, \tag{55}$$

and suppose that the eigenvalue λ is finite. Then all C_k corresponding to $s_k = -1$ are nonsingular, and setting

$$L_1 := (C_1 + \Delta C_1)^{s_1} \cdots (C_p + \Delta C_p)^{s_p} = \lambda I_m + \Delta L_1 + \tilde{L}_1 \tag{56}$$

with

$$\Delta L_1 = \sum_{k=1}^{p} (-1)^{q_k} \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) C_k^{-q_k} \Delta C_k C_k^{-q_k} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right),$$

it follows that $\|\tilde{L}_1\| = O(\|\Delta\mathbf{A}\|^2)$. For $\hat{\lambda} = \prod_{k=1}^{p} \hat{\alpha}_k^{s_k}$, applying Theorem 8 and using (55) for a given y_1 with $y_1^H x_1 \neq 0$ we have

$$\begin{aligned}
\hat{\lambda} - \lambda &= \frac{y_1^H(\Delta L_1 + \tilde{L}_1)x_1}{y_1^H x_1} + O(\|\Delta\mathbf{A}\|^2) \\
&= \frac{y_1^H \Delta L_1 x_1}{y_1^H x_1} + O(\|\Delta\mathbf{A}\|^2) \\
&= \frac{y_1^H \left\{ \sum_{k=1}^{p} (-1)^{q_k} (\prod_{j=1}^{k-1} C_j^{s_j}) C_k^{-q_k} \Delta C_k C_k^{-q_k} (\prod_{j=k+1}^{p} C_j^{s_j}) \right\} x_1}{y_1^H x_1} + O(\|\Delta\mathbf{A}\|^2).
\end{aligned}$$

By (54) and (52), we have

$$V_{k+q_k}^H \hat{U}_{k+q_k}(C_k + \Delta C_k) = \tilde{C}_k V_{k+1-q_k}^H \hat{U}_{k+1-q_k} + V_{k+q_k}^H \Delta A_k \hat{U}_{k+1-q_k}.$$

Note that $\hat{U}_k = U_k + O(\|\Delta \mathbf{A}\|^2)$. From these relations and (53), if $s_k = 1$ we get

$$\Delta C_k = H_k^{-1} V_k^H \Delta A_k U_{k+1} + C_k H_{k+1}^{-1} V_{k+1}^H \hat{U}_{k+1} - H_k^{-1} V_k^H \hat{U}_k C_k + O(\|\Delta \mathbf{A}\|^2).$$

When $s_k = -1$, then

$$-C_k^{-1} \Delta C_k C_k^{-1} = -C_k^{-1} H_{k+1}^{-1} V_{k+1}^H \left(\Delta A_k U_k C_k^{-1} + \hat{U}_{k+1} \right) - H_k^{-1} V_k^H \hat{U}_k C_k^{-1} + O(\|\Delta \mathbf{A}\|^2).$$

These two formulas have form

$$\begin{aligned}(-1)^{q_k} C_k^{-q_k} \Delta C_k C_k^{-q_k} \;&=\; (-1)^{q_k} C_k^{-q_k} H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} C_k^{-q_k} \\ &\quad + C_k^{s_k} H_{k+1}^{-1} V_{k+1}^H \hat{U}_{k+1} - H_k^{-1} V_k^H \hat{U}_k C_k^{s_k} + O(\|\Delta \mathbf{A}\|^2).\end{aligned}$$

Hence

$$\begin{aligned}&(-1)^{q_k} \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) C_k^{-q_k} \Delta C_k C_k^{-q_k} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right) \\ &= (-1)^{q_k} \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) C_k^{-q_k} H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} C_k^{-q_k} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right) \\ &\quad + \left(\prod_{j=1}^{k} C_j^{s_j} \right) H_{k+1}^{-1} V_{k+1}^H \hat{U}_{k+1} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right) \\ &\quad - \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) H_k^{-1} V_k^H \hat{U}_k \left(\prod_{j=k}^{p} C_j^{s_j} \right) + O(\|\Delta \mathbf{A}\|^2).\end{aligned}$$

Since $\prod_{k=1}^p C_k^{s_k} = \lambda I$, we have

$$\begin{aligned}&\sum_{k=1}^{p} \left\{ \left(\prod_{j=1}^{k} C_j^{s_j} \right) H_{k+1}^{-1} V_{k+1}^H \hat{U}_{k+1} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right) - \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) H_k^{-1} V_k^H \hat{U}_k \left(\prod_{j=k}^{p} C_j^{s_j} \right) \right\} \\ &\qquad = \prod_{j=1}^{p} C_j^{s_j} H_1^{-1} V_1^H \hat{U}_1 - H_1^{-1} V_1^H \hat{U}_1 \prod_{j=1}^{p} C_j^{s_j} = 0\end{aligned}$$

and hence

$$\begin{aligned}&\hat{\lambda} - \lambda \\ &= \frac{1}{y_1^H x_1} y_1^H \left\{ \sum_{k=1}^{p} (-1)^{q_k} \left(\prod_{j=1}^{k-1} C_j^{s_j} \right) C_k^{-q_k} H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} C_k^{-q_k} \left(\prod_{j=k+1}^{p} C_j^{s_j} \right) \right\} x_1 \\ &\qquad + O(\|\Delta \mathbf{A}\|^2) \\ &= \frac{1}{y_1^H x_1} y_1^H \left\{ \sum_{k=1}^{p} (-1)^{q_k} \left(\prod_{j=1}^{k+q_k-1} C_j^{s_j} \right) H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \left(\prod_{j=k+1-q_k}^{p} C_j^{s_j} \right) \right\} x_1 \\ &\qquad + O(\|\Delta \mathbf{A}\|^2). \end{aligned} \tag{57}$$

If λ is infinite then by taking the index as $(-s_1, \ldots, -s_p)$ and considering $\frac{1}{\lambda}$ we obtain a formula similar to (57).

The following theorem gives the perturbation analysis for multiple eigenvalues.

Theorem 9 *Let $\mathbf{A}$ be a regular matrix tuple with sign tuple s and let λ be an eigenvalue of $\mathbf{A}$ with multiplicity m having a complete set of eigenvectors. Let the corresponding orthonormal bases of the left and right generalized deflating subspaces $\mathbf{V}$ and $\mathbf{U}$ be chosen to satisfy (52) and (49) and let $H_k = V_k^H U_k$. Consider a perturbation $\mathbf{A} + \Delta\mathbf{A}$ with $\|\Delta\mathbf{A}\|$ sufficiently small. Then there are m associated eigenvalues of $\mathbf{A} + \Delta\mathbf{A}$, and for each such eigenvalue $\hat{\lambda}$, let $\mathbf{x} = (x_1, \ldots, x_p)$ be of unit norm satisfying (55) and let $(\hat{U}_1, \ldots, \hat{U}_p)$ satisfy (54). Then for each $l \in \{1, \ldots, p\}$ and for any y_l such that $y_l^H x_l \neq 0$, we have the following.*

a) *If λ is finite, then with $\Theta_k := \prod_{j=l}^{k+q_k-1} C_j^{s_j}$*

$$\begin{aligned}
\hat{\lambda} - \lambda &= \frac{1}{y_l^H x_l} y_l^H \left\{ \sum_{k=l}^{p+l-1} (-1)^{q_k} \Theta_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \left(\prod_{j=k+1-q_k}^{p+l-1} C_j^{s_j} \right) \right\} x_l \\
&\qquad + O(\|\Delta\mathbf{A}\|^2) \qquad (58) \\
&= \frac{1}{y_l^H x_l} \sum_{k=l}^{p+l-1} (-1)^{q_k} \left(\prod_{j=k+1-q_k}^{p+l-1} \hat{\alpha}_j^{s_j} \right) y_l^H \left\{ \Theta_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \right\} x_{k-q_k} \\
&\qquad + O(\|\Delta\mathbf{A}\|^2) \qquad (59)
\end{aligned}$$

and the following bound holds.

$$\begin{aligned}
|\hat{\lambda} - \lambda| &\leq \sum_{k=1}^{p} \left(\prod_{j=1, j\neq k}^{p} \left\| (U_{j+q_j}^H A_j U_{j+1-q_j})^{s_j} \right\| \right) \left\| (U_{k+q_k}^H A_k U_{k+1-q_k})^{-q_k} \right\|^2 \frac{\|\Delta A_k\|}{\sigma_{\min}(H_{k+q_k})} \\
&\qquad + O(\|\Delta\mathbf{A}\|^2). \qquad (60)
\end{aligned}$$

b) *If λ is infinite, then with $\Omega_k := \prod_{j=p+l-1}^{k+q_k} C_j^{-s_j}$*

$$\begin{aligned}
\frac{1}{\hat{\lambda}} &= -\frac{1}{y_l^H x_l} y_l^H \left\{ \sum_{k=l}^{p+l-1} (-1)^{q_k} \Omega_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \left(\prod_{j=k-q_k}^{l} C_j^{-s_j} \right) \right\} x_l \\
&\qquad + O(\|\Delta\mathbf{A}\|^2) \qquad (61) \\
&= -\frac{1}{y_l^H x_l} \sum_{k=l}^{p+l-1} (-1)^{q_k} \left(\prod_{j=k-q_k}^{l} \hat{\alpha}_j^{-s_j} \right) y_l^H \left\{ \Omega_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \right\} x_{k+1-q_k} \\
&\qquad + O(\|\Delta\mathbf{A}\|^2) \qquad (62)
\end{aligned}$$

and

$$\begin{aligned}
\left|\frac{1}{\hat{\lambda}}\right| &\leq \sum_{k=1}^{p} \left(\prod_{j=1, j\neq k}^{p} \left\| (U_{j+q_j}^H A_j U_{j+1-q_j})^{-s_j} \right\| \right) \left\| (U_{k+q_k}^H A_k U_{k+1-q_k})^{(q_k-1)} \right\|^2 \frac{\|\Delta A_k\|}{\sigma_{\min}(H_{k+q_k})} \\
&\qquad + O(\|\Delta\mathbf{A}\|^2). \qquad (63)
\end{aligned}$$

Proof. If $\hat{\lambda}$ is an eigenvalue of $\mathbf{C} + \Delta\mathbf{C}$, then using reordering in the periodic Schur form [9, 16]), regardless whether $\hat{\lambda}$ is simple or multiple there always exists a unit norm right eigenvector $\mathbf{x}$. Hence we have (57) if λ is finite.

Formula (57) is generated by considering the matrix product L_1 in (56). Since there exists a complete set of eigenvectors associated with λ, performing the same analysis on

$$L_l := (C_l + \Delta C_l)^{s_l} \cdots (C_{p+l-1} + \Delta C_{p+l-1})^{s_{p+l-1}},$$

for $l = 2, \ldots, p$ we get the analogous formula for $\hat{\lambda} - \lambda$. Hence we have (58). By (55) we have

$$\left(\prod_{j=k+1-q_k}^{p+l-1} C_j^{s_j}\right) x_l = \left(\prod_{j=k+1-q_k}^{p+l-1} \hat{\alpha}_j^{s_j}\right) x_{k-q_k} + z_k, \quad \|z_k\| = O(\|\!|\Delta\mathbf{A}|\!\|).$$

which implies (59). Formulae (61) and (62) are derived analogously.

The upper bounds (60) and (63) are derived by setting $y_l = x_l$ and using the fact that $C_k = U_{k+q_k}^H A_k U_{k+1-q_k}$, which is from (48). Then for λ finite, (58) yields

$$\begin{aligned}
|\hat{\lambda} - \lambda| &\le \min_l \left\| \sum_{k=l}^{p+l-1} (-1)^{q_k} \Theta_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \left(\prod_{j=k+1-q_k}^{p+l-1} C_j^{s_j}\right) \right\| + O(\|\!|\Delta\mathbf{A}|\!\|^2) \\
&\le \min_l \sum_{k=l}^{p+l-1} \left\| \Theta_k H_{k+q_k}^{-1} \right\| \left\| \prod_{j=k+1-q_k}^{p+l-1} C_j^{s_j} \right\| \|\Delta A_k\| + O(\|\!|\Delta\mathbf{A}|\!\|^2) \\
&\le \sum_{k=1}^{p} \left(\prod_{j=1, j\neq k}^{p} \left\| (U_{j+q_j}^H A_j U_{j+1-q_j})^{s_j} \right\| \right) \left\| (U_{k+q_k}^H A_k U_{k+1-q_k})^{-q_k} \right\|^2 \frac{\|\Delta A_k\|}{\sigma_{\min}(H_{k+q_k})} \\
&\qquad + O(\|\!|\Delta\mathbf{A}|\!\|^2).
\end{aligned}$$

For infinite λ we use $\prod_{k=i}^{j} C_k^{-s_k} := C_i^{-s_i} \cdots C_j^{-s_j}$ if $i \ge j$ and $\prod_{k=i}^{j} C_k^{-s_k} := I$ if $i < j$. Then (61) implies that

$$\begin{aligned}
\left|\frac{1}{\hat{\lambda}}\right| &\le \min_l \left\| \sum_{k=l}^{p+l-1} (-1)^{q_k} \Omega_k H_{k+q_k}^{-1} V_{k+q_k}^H \Delta A_k U_{k+1-q_k} \left(\prod_{j=k-q_k}^{l} C_j^{-s_j}\right) \right\| + O(\|\!|\Delta\mathbf{A}|\!\|^2) \\
&\le \min_l \sum_{k=l}^{p+l-1} \left\| \Omega_k H_{k+q_k}^{-1} \right\| \left\| \prod_{j=k-q_k}^{l} C_j^{-s_j} \right\| \|\Delta A_k\| + O(\|\!|\Delta\mathbf{A}|\!\|^2) \\
&\le \sum_{k=1}^{p} \left(\prod_{j=1, j\neq k}^{p} \left\| (U_{j+q_j}^H A_j U_{j+1-q_j})^{-s_j} \right\| \right) \left\| (U_{k+q_k}^H A_k U_{k+1-q_k})^{(q_k-1)} \right\|^2 \frac{\|\Delta A_k\|}{\sigma_{\min}(H_{k+q_k})} \\
&\qquad + O(\|\!|\Delta\mathbf{A}|\!\|^2).
\end{aligned}$$

This finishes the proof. □

The main difference between the perturbation results for simple and multiple eigenvalues is that instead of the components α_k the matrices C_k are involved. Another difference is that for $\lambda \in \{0, \infty\}$ in (44) or (45) only one ΔA_{k_0} affects the eigenvalue, while in (58) or (61) the perturbed eigenvalue seems to be influenced by all perturbations. However, by choosing a proper vector y_l in (58) or (61) it is still possible to obtain a similar result as (44) or (45). Consider for example $\lambda = 0$ in (58). Then there exists an integer l_0 such that $s_{l_0} = 1$ and C_{l_0} is singular. Let y_{l_0} be a unit norm vector such that $y_{l_0}^H C_{l_0} = 0$. Note that $s_{l_0} = 1$ and $q_{l_0} = 0$ in this case. If $y_{l_0}^H x_{l_0} \neq 0$ then equations (58) and (59) corresponding to $l = l_0$ reduce to

$$\begin{aligned}\hat{\lambda} &= \frac{y_{l_0}^H H_{l_0}^{-1} V_{l_0}^H \Delta A_{l_0} U_{l_0+1} \left(\prod_{j=l_0+1}^{p+l_0-1} C_j^{s_j}\right) x_{l_0}}{y_{l_0}^H x_{l_0}} + O(\|\Delta \mathbf{A}\|^2) \\ &= \left(\prod_{j=l_0+1}^{p+l_0-1} \hat{\alpha}_j^{s_j}\right) \frac{y_{l_0}^H H_{l_0}^{-1} V_{l_0}^H \Delta A_{l_0} U_{l_0+1} x_{l_0+1}}{y_{l_0}^H x_{l_0}} + O(\|\Delta \mathbf{A}\|^2).\end{aligned}$$

We conjecture that we can always choose such a proper y_l and similar simplified formulae hold also for all other eigenvalues.

The first order perturbation bounds for multiple eigenvalues with a complete set of eigenvectors depend on the eigenvectors of the perturbed eigenvalues which is not the case for simple eigenvalues. Since these eigenvectors are determined by the perturbation matrices, this makes the formulae less useful. However, the bounds of (60) and (63) can be used to evaluate the perturbation in the eigenvalues.

Note that even if λ has a complete set of eigenvectors, in general the matrices C_k in (49) are not diagonal if $\mathbf{U}$ is unitary. For example, if $p = 3$, $s = (1, -1, -1)$ and

$$\mathbf{A} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \right),$$

then it is easy to verify that there does not exist any triple of unitary matrices (Q_1, Q_2, Q_3), such that $Q_1^H A_1 Q_2$, $Q_3^H A_2 Q_2$ and $Q_1^H A_3 Q_3$ are simultaneously diagonal. If $p = 1$ or $p = 2$, however, C_k can be chosen to be diagonal.

Lemma 10 *Let λ be an eigenvalue of $\mathbf{A}$ with sign tuple s and suppose there exists a complete set of eigenvectors. If $p = 1$ or $p = 2$ then the orthonormal basis $\mathbf{U}$ can be chosen such that the matrices C_k, $k = 1, \ldots, p$ are all diagonal.*

Proof. For $p = 1$, with $C_1 = \lambda I$ for $s = (1)$ or $C_1 = \frac{1}{\lambda} I$ if $s = (-1)$, the result is obvious.

For $p = 2$ we only consider the case that λ is finite. The infinite case is proved analogously. Consider the case that $s = (1, -1)$, the case $s = (-1, 1)$ is analogous. We have to find unitary matrices Q_1, Q_2 such that $Q_1^H C_1 Q_2$ and $Q_1^H C_2 Q_2$ are both diagonal. Since λ is finite, C_2 must be nonsingular. Let $Q_1^H C_2 Q_2 = D_2$ be the singular value decomposition of C_2. Since $C_1 C_2^{-1} = \lambda I$, we have $Q_1^H C_1 Q_2 = \lambda D_2 =: D_1$ and the assertion follows. If $s = (1, 1)$ (or in a similar way if $s = (-1, -1)$), then we have to find unitary matrices such that $Q_1^H C_1 Q_2$ and $Q_2^H C_2 Q_1$ are both diagonal. If λ is nonzero, then let $Q_2^H C_2 Q_1 = D_2$ be the singular value decomposition of C_2. Since then D_2 must be nonsingular, using $C_1 C_2 = C_2 C_1 = \lambda I$ we have $Q_1^H C_1 Q_2 = \lambda D_2^{-1} = D_1$. If λ is zero, then let $\hat{Q}_2^H C_2 \hat{Q}_1 = \begin{bmatrix} \hat{D}_2 & 0 \\ 0 & 0 \end{bmatrix}$ be the singular value decomposition of C_2 with $\hat{D}_2$ is nonsingular. Using the commutativity, i.e., $C_1 C_2 = C_2 C_1 = 0$, it follows that the matrix $\hat{Q}_1^H C_1 \hat{Q}_2$ has the form $\begin{bmatrix} 0 & 0 \\ 0 & \hat{C}_1 \end{bmatrix}$. Let $W_1^H \hat{C}_1 W_2$ be the singular value decomposition of $\hat{C}_1$ then for $Q_1 = \hat{Q}_1 \operatorname{diag}(I, W_1)$ and $Q_2 = \hat{Q}_2 \operatorname{diag}(I, W_2)$, the matrices $Q_1^H C_1 Q_2$ and $Q_2^H C_2 Q_1$ are diagonal. $\square$

Using this lemma we obtain the classical perturbation results for matrix pencils $A - \lambda B$. considered as a formal matrix product with $p = 2$, $s = (1, -1)$ and $\mathbf{A} = (A, B)$.

Theorem 11 *Let λ be an eigenvalue of $A - \lambda B$ of multiplicity m with a complete set of eigenvectors. Let $\mathbf{U} = (U_1, U_2)$ be an orthonormal basis of the right generalized deflating subspace, such that*

$$AU_2 = U_1 C_A, \quad BU_2 = U_1 C_B,$$

with

$$C_A = \operatorname{diag}(\alpha_1, \ldots, \alpha_m), \quad C_B = \operatorname{diag}(\beta_1, \ldots, \beta_m), \quad \frac{\alpha_1}{\beta_1} = \ldots = \frac{\alpha_m}{\beta_m} = \lambda.$$

Let $\mathbf{V} = (V_1, V_2)$ be an orthonormal basis of the left generalized deflating subspace corresponding to λ and let k_0 be an integer such that $|\beta_{k_0}| = \min\{|\beta_1|, \ldots, |\beta_m|\}$ for λ finite and let $|\alpha_{k_\infty}| = \min\{|\alpha_1|, \ldots, |\alpha_m|\}$ for $\lambda = \infty$. If $\hat{A} - \lambda\hat{B} = (A + \Delta A) - \lambda(B + \Delta B)$ and $\|(\Delta A, \Delta B)\|$ is sufficiently small, then for each of the m associated eigenvalues $\hat{\lambda}$ of $\hat{A} - \lambda\hat{B}$ the following inequalities hold.

a) *If λ is nonzero and finite, then*

$$\begin{aligned} \left|\frac{\hat{\lambda} - \lambda}{\lambda}\right| &\leq \min\{\|(V_1^H U_1 C_B)^{-1} V_1^H (\tfrac{1}{\lambda}\Delta A - \Delta B) U_2\|, \\ &\qquad \|(V_1^H U_1)^{-1} V_1^H (\tfrac{1}{\lambda}\Delta A - \Delta B) U_2 C_B^{-1}\|\} + O(\|(\Delta A, \Delta B)\|^2) \\ &\leq \frac{1}{\sigma_{\min}(V_1^H U_1)} \|\frac{1}{\alpha_{k_0}}\Delta A - \frac{1}{\beta_{k_0}}\Delta B\| + O(\|(\Delta A, \Delta B)\|^2). \end{aligned}$$

b) *If $\lambda = 0$, then*

$$\begin{aligned} |\hat{\lambda}| &\leq \min\{\|(V_1^H U_1 C_B)^{-1} V_1^H \Delta A U_2\|, \|(V_1^H U_1)^{-1} V_1^H \Delta A U_2 C_B^{-1}\|\} + O(\|(\Delta A, \Delta B)\|^2) \\ &\leq \frac{1}{\beta_{k_0}\sigma_{\min}(V_1^H U_1)} \|\Delta A\| + O(\|(\Delta A, \Delta B)\|^2). \end{aligned}$$

c) *If $\lambda = \infty$, then*

$$\begin{aligned} \frac{1}{|\hat{\lambda}|} &\leq \min\{\|(V_1^H U_1 C_A)^{-1} V_1^H \Delta B U_2\|, \|(V_1^H U_1)^{-1} V_1^H \Delta B U_2 C_A^{-1}\|\} + O(\|(\Delta A, \Delta B)\|^2) \\ &\leq \frac{1}{\alpha_{k_\infty}\sigma_{\min}(V_1^H U_1)} \|\Delta B\| + O(\|(\Delta A, \Delta B)\|^2). \end{aligned}$$

Proof. If λ is finite, then using (60) and $C_A C_B^{-1} = C_B^{-1} C_A = \lambda I$, we have

$$\begin{aligned} |\hat{\lambda} - \lambda| &\leq \min\{\|(V_1^H U_1 C_B)^{-1} V_1^H (\Delta A - \lambda\Delta B) U_2\|, \\ &\qquad \|(V_1^H U_1)^{-1} V_1^H (\Delta A - \lambda\Delta B) U_2 C_B^{-1}\|\} + O(\|(\Delta A, \Delta B)\|^2). \end{aligned} \tag{64}$$

If λ is nonzero, then the first inequality is obvious, and since

$$\begin{aligned} |\hat{\lambda} - \lambda| &\leq \|C_B^{-1}\|\|(V_1^H U_1)^{-1}\|\|\Delta A - \lambda\Delta B\| + O(\|(\Delta A, \Delta B)\|^2) \\ &= \frac{1}{\sigma_{\min}(V_1^H U_1)} \|\frac{1}{\beta_{k_0}}(\Delta A - \lambda\Delta B)\| + O(\|(\Delta A, \Delta B)\|^2), \end{aligned}$$

the second inequality follows since $\alpha_{k_0}/\beta_{k_0} = \lambda$.

If $\lambda = \infty$, the assertion follows by applying the first inequality in (63). □

The bounds given here depend on $\sigma_{\min}(V_1^H U_1)$, which is the reciprocal of the condition number of λ related to the formal product AB^{-1}. One may also use the bound given by $\sigma_{\min}(V_2^H U_2)$ related to $B^{-1}A$. This can be derived as follows. Let $V_1^H A = \tilde{C}_A V_2^H$, $V_1^H B = \tilde{C}_B V_2^H$. Then $\tilde{C}_A V_2^H U_2 = V_1^H U_1 C_A$ and $\tilde{C}_B V_2^H U_2 = V_1^H U_1 C_B$. If λ is finite and C_B, $\tilde{C}_B$ are nonsingular, then the alternative bound in terms of $\sigma_{\min}(V_2^H U_2)$ follows from (64). For $\lambda = \infty$ the construction is similar. This trick can also be applied in the general case $p > 2$ if λ is simple.

Remark 5 We have already noted that for nonzero finite eigenvalues it is enough if one of the conditions (50) is satisfied. However, for zero or infinite eigenvalues, the situation is different. For example, let $p = 2$, $s = (1, 1)$ and

$$\mathbf{A} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right).$$

Then

$$B_1 = A_1 A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_2 = A_2 A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

But even though only one of the identities (50) or (51) is satisfied for some l_0, following the analysis, the perturbation results (58), (60) or (61), (63) still hold for this particular l_0. This means that even in this case we still have the first order perturbation results.

In this section we have extended the classical perturbation results for eigenvalues and eigenvectors of matrices and matrix pencils, as given, e.g., in [24], to formal products of p matrices. If the formal products consist of structured matrices, then one is interested also in structured perturbations. Typically the perturbation results change if structured perturbations are considered. This case will be studied in the next section for the special case of Hamiltonian/skew-Hamiltonian pencils.

3 Perturbation Theory for Hamiltonian/Skew-Hamiltonian Matrix Pencils

In the previous sections we have discussed the perturbation theory for formal matrix products without further assumptions on the factors A_i. These results can be used in the perturbation analysis for the periodic QZ algorithm which is used heavily in the computation of (invariant) deflating subspaces of Hamiltonian matrices [4, 5, 6] or Hamiltonian/skew-Hamiltonian pencils [2, 3].

With $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ we define the following classes of matrices. A matrix $\mathcal{H} \in \mathbb{C}^{2n \times 2n}$ is called *Hamiltonian* if $(J\mathcal{H})^H = J\mathcal{H}$ and analogously, a matrix $\mathcal{N} \in \mathbb{C}^{2n \times 2n}$ is called *skew-Hamiltonian* if $(J\mathcal{N})^H = -J\mathcal{N}$. A matrix $\mathcal{S} \in \mathbb{C}^{2n \times 2n}$ is called *symplectic* if $\mathcal{S}^H J \mathcal{S} = J$ and *unitary symplectic* if it is both unitary and symplectic. A matrix pencil $\mathcal{H} - \lambda\mathcal{N}$ with $\mathcal{H}$ Hamiltonian and $\mathcal{N}$ skew-Hamiltonian is called a *Hamiltonian/skew-Hamiltonian* pencil.

We see that Hamiltonian and skew-Hamiltonian matrices have a specific symmetry structure, and thus if we allow only structured perturbations that retain this symmetry structure, then we may expect a different perturbation analysis. For Hamiltonian matrices this analysis has recently been carried out in [17]. Using similar ideas as before for formal products of structured matrices, we can also derive the perturbation theory for Hamiltonian/skew-Hamiltonian pencils.

If a Hamiltonian/skew-Hamiltonian pencil is regular and has no purely imaginary or infinite eigenvalues, then it has been shown in [20, 21] that there exists a unitary matrix $\mathcal{Q}$ such that

$$(J\mathcal{Q}^H J^T)(\mathcal{H} - \lambda\mathcal{N})\mathcal{Q} := \mathcal{T}_{\mathcal{H}} - \lambda\mathcal{T}_{\mathcal{N}} := \begin{bmatrix} A & H \\ 0 & -A^H \end{bmatrix} - \lambda \begin{bmatrix} B & G \\ 0 & B^H \end{bmatrix}, \tag{65}$$

where $H = H^H, G = -G^H$. In many cases [3] the skew-Hamiltonian $\mathcal{N}$ is furthermore given in product form and the pencil is

$$\mathcal{H} - \lambda(J\mathcal{M}^H J^T)\mathcal{M}, \tag{66}$$

with $\mathcal{H}$ Hamiltonian. Similarly, if the pencil has no purely imaginary or infinite eigenvalues, then there exists a Hamiltonian Schur form ([23, 18]) for the Hamiltonian matrix $(J\mathcal{M}^H J^T)^{-1}\mathcal{H}\mathcal{M}^{-1}$, ($\mathcal{M}$ is nonsingular, since there is no infinite eigenvalue). Using (65) we can determine a unitary matrix $\mathcal{Q}$ and a unitary symplectic matrix $\mathcal{U}$ such that

$$\mathcal{T}_{\mathcal{H}} := J\mathcal{Q}^H J^T \mathcal{H}\mathcal{Q} = \begin{bmatrix} A & H \\ 0 & -A^H \end{bmatrix}, \quad H = H^H, \qquad \mathcal{T}_{\mathcal{M}} := \mathcal{U}^H \mathcal{M}\mathcal{Q} = \begin{bmatrix} C & F \\ 0 & D \end{bmatrix}. \tag{67}$$

The last identity implies that

$$(J\mathcal{Q}^H J^T)(J\mathcal{M}^H J^T)\mathcal{U} = J\mathcal{T}_{\mathcal{M}}^H J^T = \begin{bmatrix} D^H & -F^H \\ 0 & C^H \end{bmatrix}.$$

Combining this with (67) we get that $J\mathcal{Q}^H J^T(\mathcal{H} - \lambda J\mathcal{M}^H J^T\mathcal{M})\mathcal{Q}$ has the same block triangular form as (65).

In applications from control, see [2, 3], one is particularly interested in the perturbation theory for the eigenvalues and also the deflating subspaces spanned by the first half columns of the matrices $\mathcal{U}$ and $\mathcal{Q}$ if the perturbations are restricted to retain the matrix structure. In the following two subsections we will discuss this problem for the Hamiltonian/skew-Hamiltonian pencils and the pencils as in (66) separately.

3.1 Hamiltonian/skew-Hamiltonian pencils

The eigenvalue problem for Hamiltonian/skew-Hamiltonian pencils is a special case of the eigenvalue problem for formal products of structured matrices, with $p = 2$, $s = (1, -1)$, where $A_1 = \mathcal{H}$ and $A_2 = \mathcal{N}$. In the following we derive the structured perturbation theory for this problem.

Let $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}} = (\mathcal{H} + \Delta\mathcal{H}) - \lambda(\mathcal{N} + \Delta\mathcal{N})$ be a perturbed pencil with structured perturbations $\Delta\mathcal{H}$ Hamiltonian and $\Delta\mathcal{N}$ skew-Hamiltonian. Suppose furthermore that the original pencil $\mathcal{H} - \lambda\mathcal{N}$ has the block triangular form (65). Then we set

$$\mathcal{E}_{\mathcal{H}} = (J\mathcal{Q}^H J^T)\Delta\mathcal{H}\mathcal{Q} =: \begin{bmatrix} \Delta A & \Delta H \\ E_1 & -(\Delta A)^H \end{bmatrix}, \quad \Delta H = (\Delta H)^H, E_1 = E_1^H, \tag{68}$$

$$\mathcal{E}_{\mathcal{N}} = (J\mathcal{Q}^H J^T)\Delta\mathcal{N}\mathcal{Q} =: \begin{bmatrix} \Delta B & \Delta G \\ E_2 & (\Delta B)^H \end{bmatrix}, \quad \Delta G = -(\Delta G)^H, E_2 = -E_2^H. \tag{69}$$

Using the special transformation as in (68) and (69) the Hamiltonian and skew-Hamiltonian structures are preserved and $\mathcal{E}_{\mathcal{H}}$ and $\mathcal{E}_{\mathcal{N}}$ can be partitioned with the appropriate block structures. Partitioning $\mathcal{Q} = [Q_1, Q_2]$ with $Q_1, Q_2 \in \mathbb{C}^{2n\times n}$, we then study the perturbations in range Q_1, the right deflating subspace corresponding to the eigenvalues of $A - \lambda B$. By the definition of deflating subspaces of matrix products, the deflating subspace has the form (range $J^T Q_2$, range Q_1). As we have shown in the previous section, the perturbed unitary matrix will be $\mathcal{Q}\mathcal{Y}$, $\mathcal{Y}$ as in (12), and hence both subspaces have the same perturbation behavior. We have to determine $\mathcal{Y}_1$ and $\mathcal{Y}_2$ as in (12) to simultaneously eliminate the (2,1) blocks of $\mathcal{T}_{\mathcal{H}} + \mathcal{E}_{\mathcal{H}}$ and $\mathcal{T}_{\mathcal{N}} + \mathcal{E}_{\mathcal{N}}$. To preserve the matrix structures we require that $\mathcal{Y}_1 = J\mathcal{Y}_2 J^T$. If we set

$$\mathcal{Y}_2 = \begin{bmatrix} I_n & X^H \\ -X & I_n \end{bmatrix} \begin{bmatrix} (I_n + X^H X)^{-\frac{1}{2}} & 0 \\ 0 & (I_n + XX^H)^{-\frac{1}{2}} \end{bmatrix},$$

then the matrix X has to satisfy the quadratic equations

$$(A+\Delta A)^H X + X^H (A+\Delta A) + E_1 - X^H (H+\Delta H) X = 0, \tag{70}$$

$$(B+\Delta B)^H X - X^H (B+\Delta B) - E_2 + X^H (G+\Delta G) X = 0. \tag{71}$$

Thus, the linear transformations Φ in (7) and $\hat{\Phi}$ in (14) are replaced by the linear operators

$$\begin{aligned} \Phi_{\mathcal{H}}(X) &:= (A^H X + X^H A,\ B^H X - X^H B), \\ \hat{\Phi}_{\mathcal{H}}(X) &:= ((A+\Delta A)^H X + X^H (A+\Delta A),\ (B+\Delta B)^H X - X^H (B+\Delta B)), \end{aligned}$$

respectively. We have the following Lemma.

Lemma 12 *The following are equivalent.*

i) *The linear operator $\Phi_{\mathcal{H}}$ is nonsingular.*

ii) *The matrix pencils $A - \lambda B$ and $A^H + \lambda B^H$ have no common eigenvalues.*

iii) *The spectrum of the pencil $A - \lambda B$ does not contain purely imaginary or infinite eigenvalues, and furthermore if λ with $\operatorname{Re}\lambda \neq 0$ is in the spectrum then $-\bar{\lambda}$ is not.*

Proof. To show the equivalence of ii) and iii) observe that if λ is an eigenvalue of $A - \lambda B$, then $-\bar{\lambda}$ is an eigenvalue of $A^H + \lambda B^H$ [20]. Hence $A - \lambda B$ and $A^H + \lambda B^H$ have no common eigenvalues if and only if $A - \lambda B$ has no purely imaginary or infinite eigenvalues, and no eigenvalue pair λ, $-\bar{\lambda}$ for $\operatorname{Re}\lambda \neq 0$.

For the equivalence of i) and ii), by Lemma 1 it suffices to prove that $\Phi_{\mathcal{H}}(X)$ is nonsingular if and only if the linear transformation

$$\Phi(X,Y) = (A^H X + YA,\ B^H X - YB)$$

is nonsingular.

If $\Phi_{\mathcal{H}}(X) = 0$ has a nonzero solution X then $\Phi((X, X^H)) = 0$. Hence if Φ is nonsingular then $\Phi_{\mathcal{H}}$ is also nonsingular. If there is a nonzero (X,Y) such that $\Phi((X,Y)) = 0$, then the symmetry implies that $\Phi((Y^H, X^H)) = 0$. Hence, we either have $\Phi_{\mathcal{H}}(iX) = 0$ (if $Y = -X^H$), or $\Phi_{\mathcal{H}}(X + Y^H) = \Phi((X+Y^H, (X+Y^H)^H)) = 0$ (if $Y \neq -X^H$). In both cases $\Phi_{\mathcal{H}}$ is singular. Hence if $\Phi_{\mathcal{H}}$ is nonsingular, so is Φ. □

We can rewrite the system (70), (71) as

$$\hat{\Phi}_{\mathcal{H}}(X) + (E_1, -E_2) + \Psi_{\mathcal{H}}(X) = 0, \quad \Psi_{\mathcal{H}}(X) = (-X^H (H+\Delta H) X, X^H (G+\Delta G) X) \tag{72}$$

and then similar to Theorem 3 we have the following perturbation theorem.

Theorem 13 *If*

$$\hat{\delta}_{\mathcal{H}} := \min_{\|X\|=1} \|\hat{\Phi}_{\mathcal{H}}(X)\| > 0 \tag{73}$$

and

$$\frac{\|(E_1, E_2)\| \|(H + \Delta H, G + \Delta G)\|}{\hat{\delta}_{\mathcal{H}}^2} < \frac{1}{4}, \tag{74}$$

then (72) has a solution X which satisfies

$$\|X\| \leq \frac{2\|(E_1, E_2)\|}{\hat{\delta}_{\mathcal{H}} + \sqrt{\hat{\delta}_{\mathcal{H}}^2 - 4\|(E_1, E_2)\| \|(H + \Delta H, G + \Delta G)\|}} < 2\frac{\|(E_1, E_2)\|}{\hat{\delta}_{\mathcal{H}}}.$$

Proof. The proof is analogous to that of Theorem 3. □

Relaxing conditions (73) and (74) slightly, one obtains the following corollary.

Corollary 14 *Let*

$$\delta_{\mathcal{H}} := \min_{\|X\|=1} \|\Phi_{\mathcal{H}}(X)\|.$$

If

$$\rho_{\mathcal{H}} := \delta_{\mathcal{H}} - 2\|(\Delta A, \Delta B)\| > 0, \tag{75}$$

and

$$\frac{\|(E_1, E_2)\|(\|(H, G)\| + \|(\Delta H, \Delta G)\|)}{\rho_{\mathcal{H}}^2} < \frac{1}{4}, \tag{76}$$

then (72) has a solution X which satisfies

$$\|X\| \leq \frac{2\|(E_1, E_2)\|}{\rho_{\mathcal{H}} + \sqrt{\rho_{\mathcal{H}}^2 - 4\|(E_1, E_2)\|(\|(H, G)\| + \|(\Delta H, \Delta G)\|)}} < 2\frac{\|(E_1, E_2)\|}{\rho_{\mathcal{H}}}.$$

Using these results we obtain the following perturbation bounds for the deflating subspaces.

Theorem 15 *Let $\mathcal{H} - \lambda\mathcal{N}$ be a Hamiltonian/skew-Hamiltonian pencil that has a block upper triangular form (65). Partition $\mathcal{Q} = [Q_1, Q_2]$ with $Q_1, Q_2 \in \mathbb{C}^{2n \times n}$. Let $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}}$ be a perturbed Hamiltonian/skew-Hamiltonian pencil and let the perturbed matrices be partitioned as in (68) and (69). If conditions (73) and (74) hold, then $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}}$ has a deflating subspace* range $\hat{Q}_1$ *with $\hat{Q}_1 = \mathcal{Q}\begin{bmatrix} I_n \\ -X \end{bmatrix}(I_n + X^H X)^{-\frac{1}{2}}$, where the matrix X solves (72).*

The principal angle between range Q_1 *and* range $\hat{Q}_1$ *is less than* $\arctan(2\frac{\|(E_1, E_2)\|}{\hat{\delta}_{\mathcal{H}}})$.

Furthermore, if conditions (73), (74) are replaced by (75) and (76), respectively, then the upper bound for the principal angle is $\arctan(2\frac{\|(E_1, E_2)\|}{\rho_{\mathcal{H}}})$.

Proof. The proof is analogous to the proof for Theorem 4. □

For the perturbation of the eigenvalues we need fewer assumptions, we only assume that the pencil $\mathcal{H} - \lambda\mathcal{N}$ is regular. Let λ be an eigenvalue with algebraic multiplicity m and suppose that there exists a complete set of eigenvectors associated with λ. Since $p = 2$, let (U_1, U_2) be a corresponding orthonormal basis of the right deflating subspace with

$$\mathcal{H}U_2 = U_1 C_{\mathcal{H}}, \quad \mathcal{N}U_2 = U_1 C_{\mathcal{N}}, \tag{77}$$

where

$$C_{\mathcal{H}} = \mathrm{diag}(\alpha_1, \dots, \alpha_m), \quad C_{\mathcal{N}} = \mathrm{diag}(\beta_1, \dots, \beta_m).$$

Then we have $\lambda = \frac{\alpha_1}{\beta_1} = \ldots = \frac{\alpha_m}{\beta_m}$. The symmetry structure implies that

$$(JU_2)^H \mathcal{H} = -C_{\mathcal{H}}^H (JU_1)^H, \quad (JU_2)^H \mathcal{N} = C_{\mathcal{N}}^H (JU_1)^H,$$

and hence (JU_2, JU_1) represents the left eigenspace corresponding to the eigenvalue $-\bar{\lambda}$. Thus, if λ is purely imaginary or infinite then (JU_2, JU_1) and (U_1, U_2) are just orthonormal bases of the left and right generalized deflating subspaces. If λ is finite with $\mathrm{Re}\,\lambda \neq 0$, then let (V_1, V_2) be an orthonormal basis of the right generalized deflating subspace corresponding to $-\bar{\lambda}$ with

$$\mathcal{H}V_2 = V_1 \tilde{C}_{\mathcal{H}}, \quad \mathcal{N}V_2 = V_1 \tilde{C}_{\mathcal{N}}, \quad \Lambda(\tilde{C}_{\mathcal{H}}, \tilde{C}_{\mathcal{N}}) = \{-\bar{\lambda}\}. \tag{78}$$

Then (JV_2, JV_1) forms the left generalized deflating subspace corresponding to λ. Note that $-\bar{\lambda}$ has also multiplicity m and there again exists a complete set of the eigenvectors [21].

Using these properties and applying the results of Section 2.2 we obtain eigenvalue perturbation results for both simple and multiple eigenvalues.

Theorem 16 *Consider a regular Hamiltonian/skew-Hamiltonian pencil $\mathcal{H} - \lambda\mathcal{N}$, let λ be a simple eigenvalue and let (u_1, u_2) be the unit norm right eigenvector satisfying*

$$\mathcal{H}u_2 = \alpha_1 u_1, \quad \mathcal{N}u_2 = \alpha_2 u_1, \quad \lambda = \frac{\alpha_1}{\alpha_2}.$$

Consider the perturbed Hamiltonian/skew-Hamiltonian pencil $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}} = (\mathcal{H} + \Delta\mathcal{H}) - \lambda(\mathcal{N} + \Delta\mathcal{N})$ with $\epsilon := \|(\Delta\mathcal{H}, \Delta\mathcal{N})\|$ sufficiently small.

a) *If λ is purely imaginary or infinite then $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}}$ has unit norm eigenvectors $(\hat{u}_1, \hat{u}_2)$ satisfying $\hat{\mathcal{H}}\hat{u}_2 = \hat{\alpha}_1 \hat{u}_1$ and $\hat{\mathcal{N}}\hat{u}_2 = \hat{\alpha}_2 \hat{u}_1$, such that*

$$\hat{\alpha}_1 \alpha_2 - \hat{\alpha}_2 \alpha_1 = \frac{u_2^H (\alpha_2 J\Delta\mathcal{H} - \alpha_1 J\Delta\mathcal{N}) u_2}{u_2^H J u_1} + O(\epsilon^2).$$

b) *If $\mathrm{Re}\,\lambda \neq 0$ and (v_1, v_2) is the unit norm eigenvector corresponding to $-\bar{\lambda}$ then $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}}$ has eigenvalues $\hat{\lambda}$ and $-\bar{\hat{\lambda}}$ such that*

$$\frac{\hat{\lambda} - \lambda}{\lambda} = \frac{1}{v_2^H J u_1} v_2^H J (\frac{1}{\alpha_1}\Delta\mathcal{H} - \frac{1}{\alpha_2}\Delta\mathcal{N}) u_2 + O(\epsilon^2).$$

Proof. The proof follows directly from Theorem 6, Corollary 7, and from the symmetry property of the eigenvectors. □

Theorem 17 *Consider a regular Hamiltonian/skew-Hamiltonian pencil $\mathcal{H} - \lambda\mathcal{N}$, let λ be an eigenvalue of algebraic multiplicity m associated with a complete set of eigenvectors and let (U_1, U_2) be unitary matrices satisfying (77).*

Consider the perturbed Hamiltonian/skew-Hamiltonian pencil $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}} = (\mathcal{H} + \Delta\mathcal{H}) - \lambda(\mathcal{N} + \Delta\mathcal{N})$ and assume that $\epsilon := \|(\Delta\mathcal{H}, \Delta\mathcal{N})\|$ is sufficiently small.

If λ is purely imaginary or infinite, then for the associated eigenvalues $\hat{\lambda}$ of $\hat{\mathcal{H}} - \lambda\hat{\mathcal{N}}$ the following bounds hold.

a) *If λ is finite, then*

$$\begin{aligned} |\hat{\lambda}-\lambda| \ \leq\ & \min\{\|(U_2^H J U_1 C_{\mathcal{N}})^{-1} U_2^H J(\Delta\mathcal{H}-\lambda\Delta\mathcal{N})U_2\|, \\ & \|(U_2^H J U_1)^{-1} U_2^H J(\Delta\mathcal{H}-\lambda\Delta\mathcal{N})U_2 C_{\mathcal{N}}^{-1}\|\} + O(\epsilon^2). \end{aligned}$$

b) *If $\lambda=\infty$ then*

$$\frac{1}{|\hat{\lambda}|} \leq \min\{\|(U_2^H J U_1 C_{\mathcal{H}})^{-1} U_2^H J\Delta\mathcal{N}U_2\|, \|(U_2^H J U_1)^{-1} U_2^H J\Delta\mathcal{N}U_2 C_{\mathcal{N}}^{-1}\|\} + O(\epsilon^2).$$

If $\operatorname{Re}\lambda \neq 0$ *and* (V_1, V_2) *is unitary satisfying (78), then the associated eigenvalues* $\hat{\lambda}$ *of* $\hat{\mathcal{H}}-\lambda\hat{\mathcal{N}}$ *satisfy*

$$\begin{aligned} \left|\frac{\hat{\lambda}-\lambda}{\lambda}\right| \ \leq\ & \min\{\|(V_2^H J U_1 C_{\mathcal{N}})^{-1} V_2^H J(\frac{1}{\lambda}\Delta\mathcal{H}-\Delta\mathcal{N})U_2\|, \\ & \|(V_2^H J U_1)^{-1} V_2^H J(\frac{1}{\lambda}\Delta\mathcal{H}-\Delta\mathcal{N})U_2 C_{\mathcal{N}}^{-1}\|\} + O(\epsilon^2) \\ \leq\ & \frac{1}{\sigma_{\min}(V_2^H J U_1)}\|\frac{1}{\alpha_{k_0}}\Delta\mathcal{H}-\frac{1}{\beta_{k_0}}\Delta\mathcal{N}\| + O(\epsilon^2), \end{aligned}$$

where the integer k_0 is chosen such that $|\beta_{k_0}| = \min\{|\beta_k|,\ k=1,\ldots,p\}$.

Proof. The assertions follow from Theorem 11 and the symmetry properties of the left and right eigenvectors. □

It should be noted that if λ is purely imaginary or infinite, then the smallest singular value of the matrices $U_2^H J^T U_1$ or $U_1^H J^T U_2$ represents the reciprocal of the condition number of the eigenvalue. Moreover, $U_2^H J U_1 C_{\mathcal{N}}$ is Hermitian and $U_2^H J U_1 C_{\mathcal{H}}$ is skew-Hermitian.

3.2 The matrix pencils in (66)

We now study the matrix pencil (66) which we may consider as a formal matrix product with $p=3$, $s=(-1,1,-1)$ and $\mathbf{A}=(J\mathcal{M}^H J^T, \mathcal{H}, \mathcal{M})$. Suppose that the pencil has the form (67) and partition $\mathcal{U}=[U_1,U_2]$ and $\mathcal{Q}=[Q_1,Q_2]$ such that $U_k, Q_k \in \mathbb{C}^{2n\times n}$ for $k=1,2$. We will analyze the perturbations in $\operatorname{range} U_1, \operatorname{range} Q_1$, the generalized deflating subspace corresponding to the eigenvalues of $A-\lambda D^H C$. Let $\mathcal{H}$, $\mathcal{M}$ be perturbed to $\hat{\mathcal{H}}=\mathcal{H}+\Delta\mathcal{H}$ and $\hat{\mathcal{M}}=\mathcal{M}+\Delta\mathcal{M}$, where $\Delta\mathcal{H}$ is Hamiltonian. Set

$$(J\mathcal{Q}^H J^T)\Delta\mathcal{H}\mathcal{Q} = \begin{bmatrix} \Delta A & \Delta H \\ E_1 & -(\Delta A)^H \end{bmatrix} \tag{79}$$

and

$$\mathcal{U}^H \Delta M \mathcal{Q} = \begin{bmatrix} \Delta C & \Delta F \\ E_2 & \Delta D \end{bmatrix}. \tag{80}$$

We determine a unitary symplectic matrix

$$\mathcal{Y}_1 = \begin{bmatrix} I_n & X_1 \\ -X_1 & I_n \end{bmatrix} \begin{bmatrix} (I_n+X_1^2)^{-\frac{1}{2}} & 0 \\ 0 & (I_n+X_1^2)^{-\frac{1}{2}} \end{bmatrix}, \quad X_1 = X_1^H,$$

and a unitary matrix

$$\mathcal{Y}_2 = \begin{bmatrix} I_n & X_2^H \\ -X_2 & I_n \end{bmatrix} \begin{bmatrix} (I_n + X_2^H X_2)^{-\frac{1}{2}} & 0 \\ 0 & (I_n + X_2 X_2^H)^{-\frac{1}{2}} \end{bmatrix}$$

to eliminate the (2,1) block of $\hat{\mathcal{H}}$, $\hat{\mathcal{M}}$ and $J\hat{\mathcal{M}}^H J^T$ simultaneously. For this purpose the matrices X_1, X_2 must satisfy the quadratic equations

$$\begin{aligned} (A + \Delta A)^H X_2 + X_2^H (A + \Delta A) + E_1 - X_2^H (H + \Delta H) X_2 &= 0, & (81) \\ (D + \Delta D) X_2 - X_1 (C + \Delta C) - E_2 + X_1 (F + \Delta F) X_2 &= 0. & (82) \end{aligned}$$

Defining the linear operators

$$\begin{aligned} \Phi_{\mathcal{M}}((X_1, X_2)) &:= (A^H X_2 + X_2^H A,\ DX_2 - X_1 C), \\ \hat{\Phi}_{\mathcal{M}}((X_1, X_2)) &:= ((A + \Delta A)^H X_2 + X_2^H (A + \Delta A),\ (D + \Delta D) X_2 - X_1 (C + \Delta C)), \\ \Psi_{\mathcal{M}}((X_1, X_2)) &:= (-X_2^H (H + \Delta H) X_2,\ X_1 (F + \Delta F) X_2), \end{aligned}$$

we can rewrite the system (81), (82) as

$$\hat{\Phi}_{\mathcal{M}}((X_1, X_2)) + (E_1, -E_2) + \Psi_{\mathcal{M}}((X_1, X_2)) = 0. \tag{83}$$

We have the following lemma.

Lemma 18 *The following are equivalent.*

a) *The linear operator $\Phi_{\mathcal{M}}$ is nonsingular.*

b) *The pencils $A - \lambda D^H C$ and $A^H + \lambda C^H D$ have no common eigenvalue.*

c) *The spectrum of the pencil $A - \lambda D^H C$ does not contain purely imaginary and infinite eigenvalues, and furthermore, if λ with $\operatorname{Re}\lambda \neq 0$ is contained in the spectrum, then $-\bar{\lambda}$ is not.*

Proof. By Lemma 12 it suffices to show that $\Phi_{\mathcal{M}}$ is nonsingular if and only if

$$\Phi_{\mathcal{H}}(X) = (A^H X + X^H A,\ C^H DX - X^H D^H C)$$

is nonsingular. If $\Phi_{\mathcal{H}}$ is nonsingular, then the matrices C and D must be nonsingular, since otherwise $A - \lambda D^H C$ has an infinite eigenvalue.

If $\Phi_{\mathcal{M}}$ is singular, then there exist $X_1 (= X_1^H)$ and X_2 which are not both zero, such that

$$A^H X_2 + X_2^H A = 0 \tag{84}$$

and

$$DX_2 - X_1 C = 0. \tag{85}$$

Since $X_1 = X_1^H$, (85) implies that $C^H X_1 = X_2^H D^H$. Multiplying C^H from left to (85) we then have $C^H DX_2 - X_2^H D^H C = 0$. Combining this with (84) we get $\Phi_{\mathcal{H}}(X_2) = 0$. But since $X_1 = DX_2 C^{-1}$, it follows that $X_2 \neq 0$ and hence $\Phi_{\mathcal{H}}$ is singular, which is a contradiction. Therefore if $\Phi_{\mathcal{H}}$ is nonsingular then $\Phi_{\mathcal{M}}$ is nonsingular.

Now suppose that there exists $X \neq 0$ such that $\Phi_{\mathcal{H}}(X) = 0$. If C is nonsingular, then setting $X_1 = DXC^{-1}$ and $X_2 = X$, we have $X_1 = X_1^H$ and $\Phi_{\mathcal{M}}((X_1, X_2)) = 0$. If C is singular, then let $C = U \begin{bmatrix} \Gamma & 0 \\ 0 & 0 \end{bmatrix} V^H$, with U, V unitary, and Γ nonsingular be the singular value decomposition of C. Then with $X_1 = U \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} U^H$ and $X_2 = 0$, we have $\Phi_{\mathcal{M}}((X_1, X_2)) = 0$. Hence, if $\Phi_{\mathcal{H}}$ is singular then $\Phi_{\mathcal{M}}$ is singular. □

We obtain the following perturbation bounds.

Theorem 19 *If*

$$\hat{\delta}_{\mathcal{M}} := \min_{|||(X_1,X_2)|||=1} |||\hat{\Phi}_{\mathcal{M}}((X_1, X_2))||| > 0 \tag{86}$$

and

$$\frac{|||(E_1, E_2)||| \, |||(H + \Delta H, F + \Delta F)|||}{\hat{\delta}_{\mathcal{M}}^2} < \frac{1}{4}, \tag{87}$$

then (83) has a solution (X_1, X_2) *which satisfies*

$$|||(X_1, X_2)||| \leq \frac{2|||(E_1, E_2)|||}{\hat{\delta}_{\mathcal{M}} + \sqrt{\hat{\delta}_{\mathcal{M}}^2 - 4|||(E_1, E_2)||| \, |||(H + \Delta H, F + \Delta F)|||}} < 2\frac{|||(E_1, E_2)|||}{\hat{\delta}_{\mathcal{M}}}.$$

Proof. The proof is analogous to the proof of Theorem 3. □

Under slightly stronger assumptions we have the following corollary.

Corollary 20 *Let*

$$\delta_{\mathcal{M}} := \min_{|||(X_1,X_2)|||=1} |||\Phi_{\mathcal{M}}((X_1, X_2))|||.$$

If

$$\rho_{\mathcal{M}} := \delta_{\mathcal{M}} - \max\{2\|\Delta A\|, |||(\Delta C, \Delta D)|||\} > 0, \tag{88}$$

and

$$\frac{|||(E_1, E_2)|||(|||(H, F)||| + |||(\Delta H, \Delta F)|||)}{\rho_{\mathcal{M}}^2} < \frac{1}{4}, \tag{89}$$

then (83) has a solution (X_1, X_2) *which satisfies*

$$|||(X_1, X_2)||| \leq \frac{2|||(E_1, E_2)|||}{\rho_{\mathcal{M}} + \sqrt{\rho_{\mathcal{M}}^2 - 4|||(E_1, E_2)|||(|||(H, F)||| + |||(\Delta H, \Delta F)|||)}} < 2\frac{|||(E_1, E_2)|||}{\rho_{\mathcal{M}}}.$$

We then finally have the perturbation result for the generalized deflating subspace.

Theorem 21 *Let* $\mathcal{H} - \lambda(J\mathcal{M}^H J^T)\mathcal{M}$ *be a Hamiltonian/skew-Hamiltonian pencil in the block upper triangular form (67) and let* $\mathcal{Q} = [Q_1, Q_2]$, $\mathcal{U} = [U_1, U_2]$ *with* $Q_1, Q_2, U_1, U_2 \in \mathbb{C}^{2n \times n}$.

Let the perturbed matrices $\hat{\mathcal{H}}$, $\hat{\mathcal{M}}$ *be partitioned as in (79) and (80). If conditions (86) and (87) hold, then* $\hat{\mathcal{H}} - \lambda(J\hat{\mathcal{M}}^H J^T)\hat{\mathcal{M}}$ *has a generalized deflating subspace given by* range $\hat{U}_1$, range $\hat{Q}_1$ *with*

$$\hat{U}_1 = \mathcal{U} \begin{bmatrix} I_n \\ -X_1 \end{bmatrix} (I_n + X_1^2)^{-\frac{1}{2}}, \quad \hat{Q}_1 = \mathcal{Q} \begin{bmatrix} I_n \\ -X_2 \end{bmatrix} (I_n + X_2^H X_2)^{-\frac{1}{2}},$$

where the matrix pair (X_1, X_2) *solves (83).*

An upper bound for the largest principal angle between range U_1 *and* range $\hat{U}_1$ *or between* range Q_1 *and* range $\hat{Q}_1$, *respectively, is given by* $\arctan(2\frac{|||(E_1,E_2)|||}{\hat{\delta}_{\mathcal{M}}})$.

If conditions (86) and (87) are replaced by (88) and (89), then the upper bound for the largest principal angle is $\arctan(2\frac{|||(E_1,E_2)|||}{\rho_{\mathcal{M}}})$.

Proof. The proof is analogous to the proof for Theorem 4. □

For the perturbations in the eigenvalues there are still further special properties that follow from the matrix structures. Let $\mathcal{H} - \lambda(J\mathcal{M}^H J^T)\mathcal{M}$ be regular and let λ be an eigenvalue with multiplicity m having a complete set of eigenvectors. Let $\mathbf{U} = (U_1, U_2, U_3)$ be unitary such that

$$J\mathcal{M}^H J^T U_1 = U_2 C_1, \quad \mathcal{H} U_3 = U_2 C_2, \quad \mathcal{M} U_3 = U_1 C_3, \tag{90}$$

and

$$C_1^{-1} C_2 C_3^{-1} = C_2 C_3^{-1} C_1^{-1} = C_3^{-1} C_2 C_1^{-1} = \lambda I_m.$$

Then using the matrix structure, if λ is purely imaginary or infinite then (JU_1, JU_3, JU_2) is an orthonormal basis of the left generalized deflating subspace corresponding to λ. Moreover, from (90) we have

$$\begin{aligned} C_1^H (U_2^H J U_3) &= (U_1^H J U_1) C_3, \\ C_3^H (U_1^H J U_1) &= (U_3^H J U_2) C_1, \\ C_2^H (U_2^H J U_3) &= -(U_3^H J U_2) C_2. \end{aligned} \tag{91}$$

If $\operatorname{Re}\lambda \neq 0$ and (V_1, V_2, V_3) represents an orthonormal basis of the right generalized deflating subspace corresponding to $-\bar{\lambda}$, i.e.,

$$J\mathcal{M}^H J^T V_1 = V_2 \tilde{C}_1, \quad \mathcal{H} V_3 = V_2 \tilde{C}_2, \quad \mathcal{M} V_3 = V_1 \tilde{C}_3, \tag{92}$$

and

$$\tilde{C}_1^{-1} \tilde{C}_2 \tilde{C}_3^{-1} = \tilde{C}_2 \tilde{C}_3^{-1} \tilde{C}_1^{-1} = \tilde{C}_3^{-1} \tilde{C}_2 \tilde{C}_1^{-1} = -\bar{\lambda} I_m,$$

then (JV_1, JV_3, JV_2) represents an orthonormal basis of the left generalized deflating subspace corresponding to λ. Similarly

$$\begin{aligned} \tilde{C}_1^H (V_2^H J U_3) &= (V_1^H J U_1) C_3, \\ \tilde{C}_3^H (V_1^H J U_1) &= (V_3^H J U_2) C_1, \\ \tilde{C}_2^H (V_2^H J U_3) &= -(V_3^H J U_2) C_2. \end{aligned} \tag{93}$$

Using these properties we have the following perturbation results for simple and multiple eigenvalues.

Theorem 22 *Let $\mathcal{H} - \lambda(J\mathcal{M}^H J^T)\mathcal{M}$ be a regular Hamiltonian/skew-Hamiltonian pencil and let λ be a simple eigenvalue. Let (u_1, u_2, u_3) be the unit norm right eigenvector satisfying*

$$(J\mathcal{M}^H J^T) u_1 = \alpha_1 u_2, \quad \mathcal{H} u_3 = \alpha_2 u_2, \quad \mathcal{M} u_3 = \alpha_3 u_1, \quad \frac{\alpha_2}{\alpha_1 \alpha_3} = \lambda,$$

and let $\hat{\mathcal{H}} = \mathcal{H} + \Delta\mathcal{H}$, $\hat{\mathcal{M}} = \mathcal{M} + \Delta\mathcal{M}$ with $\Delta\mathcal{H}$ Hamiltonian. Furthermore, let $\epsilon := \|(\Delta\mathcal{H}, \Delta\mathcal{M})\|$ be sufficiently small. Then $\hat{\mathcal{H}} - \lambda(J\hat{\mathcal{M}}^H J^T)\hat{\mathcal{M}}$ has the unit norm eigenvectors $\hat{u}_1, \hat{u}_2, \hat{u}_3$ satisfying

$$J\hat{\mathcal{M}}^H J^T \hat{u}_1 = \hat{\alpha}_1 \hat{u}_2, \quad \hat{\mathcal{H}} \hat{u}_3 = \hat{\alpha}_2 \hat{u}_2, \quad \hat{\mathcal{M}} \hat{u}_3 = \hat{\alpha}_3 \hat{u}_1, \quad \hat{\lambda} = \frac{\hat{\alpha}_2}{\hat{\alpha}_1 \hat{\alpha}_3}.$$

a) *If λ is purely imaginary, then*

$$\hat{\lambda} - \lambda = \frac{u_3^H J \Delta\mathcal{H} u_3}{\alpha_1 \alpha_3 u_3^H J u_2} - \lambda \left(\frac{u_3^H (\Delta\mathcal{M})^H J u_1}{\alpha_1 u_3^H J u_2} - \frac{u_1^H J \Delta\mathcal{M} u_3}{\alpha_3 u_1^H J u_1} \right) + O(\epsilon^2).$$

b) *If* $\lambda = \infty$, *then*

$$\frac{1}{\hat{\lambda}} = O(\epsilon^2).$$

c) *If* $\operatorname{Re}\lambda \neq 0$ *and* (v_1, v_2, v_3) *is a unit norm left eigenvector corresponding to* $-\bar{\lambda}$, *then* $\hat{\mathcal{H}} - \lambda(J\hat{\mathcal{M}}^H J^T)\hat{\mathcal{M}}$ *has eigenvalues* $\hat{\lambda}$ *and* $-\bar{\hat{\lambda}}$, *such that*

$$\frac{\hat{\lambda}-\lambda}{\lambda} = \frac{v_3^H J\Delta\mathcal{H}u_3}{\alpha_2 v_3^H J u_2} - \frac{v_3^H(\Delta\mathcal{M})^H J u_1}{\alpha_1 v_3^H J u_2} - \frac{v_1^H J\Delta\mathcal{M}u_3}{\alpha_3 v_1^H J u_1} + O(\epsilon^2).$$

Proof. Consider the formal product with $p = 3$, $s_1 = s_3 = -1$, $s_2 = 1$, and factors $A_1 = J\mathcal{M}^H J^T$, $A_2 = \mathcal{H}$ and $A_3 = \mathcal{M}$. Consider perturbations $\Delta A_1 = J\Delta\mathcal{M}^H J^T$, $\Delta A_2 = \Delta\mathcal{H}$ and $\Delta A_3 = \Delta\mathcal{M}$. If λ is finite, then the results follows from Corollary 7.

If $\lambda = \infty$, then by (91) we have $\bar{\alpha}_1 u_2^H J u_3 = \alpha_3 u_1^H J u_1$. Since $\alpha_1\alpha_3 = 0$, $u_2^H J u_3 \neq 0$ and $u_1^H J u_1 \neq 0$ we have $\alpha_1 = \alpha_3 = 0$. Hence from Corollary 7 we get $1/\hat{\lambda} = O(\epsilon^2)$. □

For multiple eigenvalues the result is as follows.

Theorem 23 *Let* $\mathcal{H} - \lambda(J\mathcal{M}^H J^T)\mathcal{M}$ *be a regular Hamiltonian/skew-Hamiltonian pencil and let* λ *be an eigenvalue of algebraic multiplicity* m *with a complete set of eigenvectors. Let* (U_1, U_2, U_3) *be unitary satisfying (90) and consider perturbed matrices* $\hat{\mathcal{H}} = \mathcal{H} + \Delta\mathcal{H}$ *with* $\hat{\mathcal{M}} = \mathcal{M} + \Delta\mathcal{M}$ *with* $\Delta\mathcal{H}$ *Hamiltonian and* $\epsilon := \|(\Delta\mathcal{H}, \Delta\mathcal{M})\|$ *sufficiently small. Then for the associated eigenvalues* $\hat{\lambda}$ *of the perturbed problem we obtain the following bounds.*

a) *If* λ *is purely imaginary then*

$$\begin{aligned}|\hat{\lambda}-\lambda| \leq \min\{&\|(U_1^H J U_1)^{-1} C_3^{-H} E_a C_3^{-1}\|,\\ &\|(U_3^H J U_2)^{-1} E_a C_3^{-1} C_1^{-1}\|, \|(U_2^H J U_3)^{-1} C_1^{-H} C_3^{-H} E_a\|\} + O(\epsilon^2),\end{aligned}$$

where $E_a = \lambda(U_3^H(\Delta\mathcal{M})^H J U_1 C_3 + C_3^H U_1^H J\Delta\mathcal{M}U_3) - U_3^H J\Delta\mathcal{H}U_3$.

b) *If* $\lambda = \infty$, *then*

$$\frac{1}{|\hat{\lambda}|} \leq \min\{\|(U_1^H J U_1)^{-1} E_\infty\|, \|(U_3^H J U_2)^{-1} E_b C_2^{-1}\|, \|C_2^{-1}(U_3^H J U_2)^{-1} E_b\|\} + O(\epsilon^2),$$

where

$$E_\infty = C_1^H C_2^{-H} U_3^H \Delta\mathcal{M}^H J U_1 - U_1^H J\Delta\mathcal{M}U_3 C_2^{-1} C_1 - C_1^H C_2^{-H} U_3^H J\Delta\mathcal{H}U_3 C_2^{-1} C_1$$

and $E_b = U_3^H(\Delta\mathcal{M})^H J U_1 C_3 + C_3^H U_1^H J\Delta\mathcal{M}U_3$.

c) *If* $\operatorname{Re}\lambda \neq 0$ *and* (V_1, V_2, V_3) *represents an orthonormal basis of the right generalized deflating subspace corresponding to* $-\bar{\lambda}$ *satisfying (92), then there are* m *eigenvalues* $\hat{\lambda}$ *of* $\hat{\mathcal{H}} - \lambda(J\hat{\mathcal{M}}^H J^T)\hat{\mathcal{M}}$ *that satisfy*

$$\begin{aligned}\left|\frac{\hat{\lambda}-\lambda}{\lambda}\right| \leq \min\{&\|(V_1^H J U_1)^{-1} \tilde{C}_3^{-H} E_c C_3^{-1}\|,\\ &\|(V_3^H J U_2)^{-1} E_c C_3^{-1} C_1^{-1}\|, \|(V_2^H J U_3)^{-1} \tilde{C}_1^{-H} \tilde{C}_3^{-H} E_c\|\} + O(\epsilon^2),\end{aligned}$$

where $E_c = V_3^H(\Delta\mathcal{M})^H J U_1 C_3 + \tilde{C}_3^H U_1^H J\Delta\mathcal{M}U_3 - \frac{1}{\lambda} V_3^H J\Delta\mathcal{H}U_3$.

Proof. If λ is purely imaginary the result follows from (58) of Theorem 9 and the properties of (91) and (90). If $\lambda = \infty$ then the bound follows from (61), (91) and the fact that $C_1C_2^{-1}C_3 = C_2^{-1}C_3C_1 = C_3C_1C_2^{-1} = 0$. If $\operatorname{Re}\lambda \neq 0$ the bound again follows from (58), (93) and (92). □

Note that in Theorem 23 the matrix E_a has skew-Hermitian and Hermitian parts which are composed by $\Delta\mathcal{M}$ and $\Delta\mathcal{H}$, respectively. Furthermore E_∞ is Hermitian and E_b is skew-Hermitian.

4 Conclusion

We have analyzed the perturbation theory for generalized deflating subspaces and eigenvalues of a formal matrix product. The perturbation bounds can be used to estimate the errors of the generalized deflating subspaces and eigenvalues when they are computed by the periodic QR or QZ algorithm. As an application we have studied the perturbation theory for Hamiltonian/skew-Hamiltonian pencils. The symmetry structure of the matrices then leads to a symmetry structure in the perturbation results and hence sharper perturbation bounds. Although we have presented all results for complex matrices, it should be noted that similar results hold for real pencils.

References

[1] P. Benner and R. Byers. Evaluating products of matrix pencils and collapsing matrix products. Mathematics Research Reports 99-12-01, Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA, December 1999. Available from `http://www.math.ukans.edu/reports/1999.html`.

[2] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of embedded Hamiltonian pencils. Technical Report SFB393/99-15, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, Germany, June 1999. Available from `http://www.tu-chemnitz.de/sfb393/sfb99pr.html`.

[3] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical methods for linear-quadratic and H_∞ control problems. In G. Picci and D.S. Gilliam, editors, *Dynamical Systems, Control, Coding, Computer Vision: New Trends, Interfaces, and Interplay*, volume 25 of *Progress in Systems and Control Theory*, pages 203–222. Birkhäuser, Basel, 1999.

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

[5] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[6] P. Benner, V. Mehrmann, and H. Xu. A note on the numerical solution of complex Hamiltonian and skew-Hamiltonian eigenvalue problems. *Electr. Trans. Num. Anal.*, 8:115–126, 1999.

[7] S. Bittanti and P. Colaneri. Analysis of the periodic Lyapunov and Riccati equations via canonical decomposition. *SIAM J. Cont. Optim.*, 24(6):1138–1149, 1986.

[8] S. Bittanti, P. Colaneri, and G. De Nicolao. The periodic Riccati equation. In S. Bittanti, A.J. Laub, and J.C. Willems, editors, *The Riccati Equation*, pages 127–162. Springer-Verlag, Berlin, Heidelberg, Germany, 1991.

[9] A. Bojanczyk, G.H. Golub, and P. Van Dooren. The periodic Schur decomposition. algorithms and applications. In F.T. Luk, editor, *Advanced Signal Processing Algorithms, Architectures, and Implementations III*, volume 1770 of *Proc. SPIE*, pages 31–42, 1992.

[10] R. Byers, C. He, and V. Mehrmann. Where is the nearest non-regular pencil? *Linear Algebra Appl.*, 285:81–105, 1998.

[11] J.W. Demmel and B. Kågström. The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. I: theory and algorithms. *ACM Trans. Math. Software*, 19(2):160–174, 1993.

[12] J.W. Demmel and B. Kågström. The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. II: software and applications. *ACM Trans. Math. Software*, 19(2):175–201, 1993.

[13] J. Francis. The QR–transformation, part I, II. *Comput. J.*, 4:265–271, 332–345, 1961.

[14] F.R. Gantmacher. *Theory of Matrices*, volume 2. Chelsea, New York, 1959.

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

[16] J.J. Hench and A.J. Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39:1197–1210, 1994.

[17] M.M. Konstantinov, V. Mehrmann, and P.Hr. Petkov. Perturbation analysis for the Hamiltonian Schur form. *SIAM J. Matrix Anal. Appl.*, to appear. *See also*: Preprint 98-17, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 1998.

[18] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302–303:469–533, 1999.

[19] D.G. Luenberger. Boundary recursion for descriptor variable system. *IEEE Trans. Automat. Control*, AC-34:287–292, 1989.

[20] C. Mehl. *Compatible Lie and Jordan algebras and applications to structured matrices and pencils.* Dissertation, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz (FRG), 1998.

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

[22] C.B. Moler and G.W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973.

[23] C.C. Paige and C.F. Van Loan. A Schur decomposition for Hamiltonian matrices. *Linear Algebra Appl.*, 14:11–32, 1981.

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

[25] P. Van Dooren. Two point boundary value and periodic eigenvalue problems. In O. Gonzalez, editor, *Proc. 1999 IEEE Intl. Symp. CACSD, Kohala Coast-Island of Hawai'i, Hawai'i, USA, August 22–27, 1999* (CD-Rom), pages 58–63, 1999.

[26] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford University Press, Oxford, 1965.

Berichte aus der Technomathematik

ISSN 1435-7968

http://www.math.uni-bremen.de/zetem/berichte.html

— Vertrieb durch den Autor —

Reports

Stand: 14. Februar 2000

98–01. Peter Benner, Heike Faßbender:
An Implicitly Restarted Symplectic Lanczos Method for the Symplectic Eigenvalue Problem, Juli 1998.

98–02. Heike Faßbender:
Sliding Window Schemes for Discrete Least-Squares Approximation by Trigonometric Polynomials, Juli 1998.

98–03. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Parallel Partial Stabilizing Algorithms for Large Linear Control Systems, Juli 1998.

98–04. Peter Benner:
Computational Methods for Linear–Quadratic Optimization, August 1998.

98–05. Peter Benner, Ralph Byers, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, August 1998.

98–06. Lars Grüne, Fabian Wirth:
On the rate of convergence of infinite horizon discounted optimal value functions, November 1998.

98–07. Peter Benner, Volker Mehrmann, Hongguo Xu:
A Note on the Numerical Solution of Complex Hamiltonian and Skew-Hamiltonian Eigenvalue Problems, November 1998.

98–08. Eberhard Bänsch, Burkhard Höhn:
Numerical simulation of a silicon floating zone with a free capillary surface, Dezember 1998.

99–01. Heike Faßbender:
The Parameterized SR Algorithm for Symplectic (Butterfly) Matrices, Februar 1999.

99–02. Heike Faßbender:
Error Analysis of the symplectic Lanczos Method for the symplectic Eigenvalue Problem, März 1999.

99–03. Eberhard Bänsch, Alfred Schmidt:
Simulation of dendritic crystal growth with thermal convection, März 1999.

99–04. Eberhard Bänsch:
Finite element discretization of the Navier-Stokes equations with a free capillary surface, März 1999.

99–05. Peter Benner:
Mathematik in der Berufspraxis, Juli 1999.

99–06. Andrew D.B. Paice, Fabian R. Wirth:
Robustness of nonlinear systems and their domains of attraction, August 1999.

99–07. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Balanced Truncation Model Reduction of Large-Scale Dense Systems on Parallel Computers, September 1999.

99–08. Ronald Stöver:
Collocation methods for solving linear differential-algebraic boundary value problems, September 1999.

99–09. Huseyin Akcay:
Modelling with Orthonormal Basis Functions, September 1999.

99–10. Heike Faßbender, D. Steven Mackey, Niloufer Mackey:
Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems, Oktober 1999.

99–11. Peter Benner, Vincente Hernández, Antonio Pastor:
On the Kleinman Iteration for Nonstabilizable System, Oktober 1999.

99–12. Peter Benner, Heike Faßbender:
A Hybrid Method for the Numerical Solution of Discrete-Time Algebraic Riccati Equations, November 1999.

99–13. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Numerical Solution of Schur Stable Linear Matrix Equations on Multicomputers, November 1999.

99–14. Eberhard Bänsch, Karol Mikula:
Adaptivity in 3D Image Processing, Dezember 1999.

00–01. Peter Benner, Volker Mehrmann, Hongguo Xu:
Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices, Januar 2000.