Perturbation Results Related to Palindromic Eigenvalue Problems

E. K.-W. Chu* W.-W. Lin† C.-S. Wang‡

Abstract

We investigate the perturbation of the palindromic eigenvalue problem for the matrix quadratic $P(\lambda) \equiv \lambda^2 A_1^\star + \lambda A_0 + A_1$, with A_0, $A_1 \in \mathcal{C}^{n\times n}$ and $A_0^\star = A_0$ ($\star = T, H$). The perturbation of eigenvalues, in terms of general matrix polynomials, palindromic pencils, (semi-Schur) anti-triangular canonical forms and differentiation, are discussed.

Keywords: Anti-triangular form, eigenvalue, eigenvector, matrix polynomial, palindromic eigenvalue problem, palindromic linearization, palindromic pencil, perturbation

1 Introduction

Consider the matrix quadratic

$$P(\lambda) \equiv \lambda^2 A_1^\star + \lambda A_0 + A_1$$

where A_0, $A_1 \in \mathcal{C}^{n\times n}$ and $A_0^\star = A_0$ ($\star = T, H$), and the corresponding *palindromic* quadratic eigenvalue problem

$$P(\lambda)x = 0\ , \quad x \neq 0\ . \tag{1}$$

In this paper, we shall consider only *regular* matrix polynomials $P(\lambda)$, in the sense that $\det P(\lambda) \not\equiv 0$.

*School of Mathematical Sciences, Building 28, Monash University, VIC 3800, AUSTRALIA. mailto:eric.chu@sci.monash.edu.au

†Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, TAIWAN. mailto:wwlin@am.nctu.edu.tw

‡Department of Mathematics, National Cheng Kung University, Tainan 701, TAIWAN. mailto:cswang@math.ncku.edu.tw

From the transpose or Hermitian of (1), a palindromic eigenvalue problem possesses a spectrum $\sigma(P)$ containing both λ and its "reciprocal" $1/\lambda^\star$ (with 0 and ∞ considered to be reciprocal to each other). Under favourable conditions, the eigenvalue problem of the original matrix polynomial $P(\lambda)$ has a palindromic linearization [6, 10, 16] of the form $\lambda Z \pm Z^\star$.

We can transform $\lambda Z - Z^\star$ to the form $\nu(-Z) + (-Z)^\star$ with $\nu = -\lambda$. Similarly, $\lambda^2 A_1^\star + \lambda A_0 + A_1$ and $\nu^2 A_1^\star - \nu A_0 + A_1$ define equivalent palindromic eigenvalue problems. Consequently, and for simplicity of presentation, we shall concentrate on the palindromic eigenvalue problem. Anti-palindromic or odd/even eigenvalue problems [16, 17, 18] can be treated similarly.

A great foundation for the solution of palindromic eigenvalue problems has been laid by Hilliges, Mackey, Mehl and Mehrmann (see [13, 16, 17] by various members of the group). An alternative approach in tackling the problem with structure-preserving doubling algorithms can be found in [6]. There has been much recent interest in quadratic eigenvalue problems [22]. An important example of palindromic eigenvalue problems can be found in the vibration analysis of fast trains; see [13, 14] for general introductions and [12] for details. For general perturbation of eigenvalues for polynomial eigenvalue problems, see [5, 21]; consult also [8, 15] for some related results.

This paper is organized as follows. In Sections 2–5, the perturbation of eigenvalues, in terms of general matrix polynomials, palindromic linearizations, the anti-triangular canonical form [16, 17, 18] and the semi-Schur anti-triangular canonical form, is discussed using the Bauer-Fike technique for perturbations of arbitrary size [5]. The derivatives of eigenvalues and eigenvectors of $P(\lambda)$ are considered in Section 6. The paper is concluded in Section 7. The results in Section 3–5 are applicable for the more general palindromic pencils $Z - \lambda Z^\star$, which may not be a linearization of any matrix polynomials.

Note that Sun and Stewart's implicit function approach [20] has been applied to palindromic linearizations, general matrix quadratics and palindromic eigenvalue problems in [7], to obtain perturbation results for (simple) eigenvalues and the corresponding deflating subspaces.

It is important to distinguish between the different perturbation techniques. Note that the Bauer-Fike technique allows perturbations of arbitrary size and the clustering of eigenvalues (thus the corresponding deflating subspaces) may vary greatly. Consequently, it is meaningless to talk about the perturbation of eigenvectors or deflating subspaces in Bauer-Fike type perturbation theorems in Sections 2–5. Perturbation results from differentiation or implicit function approaches are valid only for asymptotically small perturbations but results for deflating subspaces are available. Notice that only simple eigenvalues (or sum and averages of multiple eigenvalues) are differentiable, and generalized derivatives (or subgradients) have to be utilized in general [4]. Note that the naive differentiation technique in Section 6 assumes differentiability, which can only be proved rigorously using tools like the implicit function theorem in Section 7.

Next, we want to present a few words of warning. Comparison of perturbation results is a risky art. Typically, error bounds and condition numbers are simplified upper bounds of more complicated quantities and a better (worse) upper bound does not always imply a smaller (bigger) error. Furthermore, optimization of such upper bounds is often possible but seldom attempted because of costs or inconvenience, making such comparison of perturbation results even more perilous. Consequently, we do not claim to have found the "best" perturbation results, if they exist. Quite often, perturbation results are interpreted qualitatively rather than applied quantitatively, indicating when things may go wrong or pitfalls to avoid. Nevertheless, our collection of perturbation results are the only one presently available for palindromic eigenvalue problems, in addition to those in [2, 7, 15], and should be useful in the associated investigations. Lastly, conditions qualifying when perturbations are large or (asymptotically) small may be written down but are complicated and rarely checked. Again, such perturbation results may have to be used qualitatively rather than quantitatively.

2 Bauer-Fike Theorem for General Matrix Polynomials

The unstructured perturbation result in Theorem 2.1 below, for general matrix polynomials, may not be directly applicable or satisfactory for palindromic eigenvalue problems. However, it serves as a reference for the structured perturbation results in later Sections. Also, palindromic eigenvalue problems are sometimes perturbed in an unstructured manner. One example is when the QZ algorithm [11] is applied to an associated palindromic linearization $Z - \lambda Z^{\star}$. The associated perturbation problem has to be treated as an unstructured one, using the theorem below.

We quote without proof [5, Theorem 4.2] on the perturbation of eigenvalues of a general matrix polynomial:

Theorem 2.1 *Consider a regular matrix polynomial*

$$L(\alpha, \beta) \equiv \sum_{j=0}^{l} B_j \alpha^j \beta^{l-j} \ ,$$

and its perturbation

$$\widetilde{L}(\alpha,\beta) \equiv \sum_{j=0}^{l} \widetilde{B}_j \alpha^j \beta^{l-j} \ , \quad \widetilde{B}_j \equiv B_j + \delta B_j \ (j=0,\cdots,l) \ .$$

Let (X,T,Z) *be a resolvent triple [5, 10] for* L *constructed using some finite and infinite Jordan pairs* J_F *and* J_∞*. For* $(\alpha_i,\beta_i) \in \sigma(L)$ *and* $(\alpha,\beta) \in \sigma(\widetilde{L})$*, with the scaling* $|\alpha_i|^2 + |\beta_i|^2 = 1 = |\alpha|^2 + |\beta|^2$*, the spectral variation of* $\widetilde{L}$ *from* L *is defined as*

$$s_L(\widetilde{L}) \equiv \max_{(\alpha,\beta)}\{s_{(\alpha,\beta)}\} \ , \quad s_{(\alpha,\beta)} \equiv \min_i \{|\alpha\beta_i - \beta\alpha_i|\} \ .$$

Let p *be the maximum dimension of the Jordan blocks in* J_F *or* J_∞*.*

Then for $\|\cdot\| = \|\cdot\|_\tau$ $(\tau = 1,2,\infty)$*, we have*

$$s_{(\alpha,\beta)} \leq \max\{\theta_1, \theta_1^{1/p}\} \ , \quad \theta_1 \equiv p\mathcal{F}\kappa\Delta \ , \tag{2}$$

where $c_1 = 1$ *(for* $\tau = 1,2$*) or* $\sqrt{l}$ *(for* $\tau = \infty$*), and*

$$\mathcal{F} \equiv c_1\sqrt{\frac{l+1}{2}} \ , \quad \kappa \equiv \|X\| \cdot \|Z\| \ , \quad \Delta \equiv \|[\delta B_0, \cdots, \delta B_l]\| \ .$$

Also, we have

$$s_L(\widetilde{L}) \leq \max\{\theta_1, \theta_1^{1/p}\} \ . \tag{3}$$

Comments:

(a) Note that we use the representation (α,β) for $\lambda = \alpha/\beta$ in the above theorem.

(b) In the palindromic case, we have $l = 2$ and $B_0 = A_1^\star = B_2^\star$ and $B_1 = A_0 = A_0^\star$. Ultimately, the perturbation of the palindromic eigenvalues is controlled by θ_1 in (3), which is in turn dominated by the error term involving $\|[\delta A_0, \delta A_1]\|$. The condition of the eigenvalues will be poor when κ is large, or when deflating subspaces for different eigenvalues are getting "close" to each other, making the resolution of the spectrum more and more difficult. Note also that the perturbation in δA_0 may be nonsymmetric, pushing a pair of reciprocal palindromic eigenvalues to one which are not reciprocal (or approximately so when δA_0 is small). For a symmetric δA_0, we only have to consider the perturbation of half of the eigenvalues, because of the palindromic structure.

(c) Based on Theorem 2.1, we can consider the perturbation of a cluster of eigenvalues; for details, see [5, Section 5.2]. A cluster (defined later in (4)) can be one simple eigenvalue, a group of multiple eigenvalues or a group of neighbouring eigenvalues. For $(\alpha, \beta) \notin \sigma(L)$, assume that the decomposition of the resolvent

$$L(\alpha, \beta)^{-1} = X_1 T_1(\alpha, \beta)^{-1} Z_1 + X_2 T_2(\alpha, \beta)^{-1} Z_2$$

where $([X_1, X_2], T_1 \oplus T_2, [Z_1, Z_2])$ is a resolvent triple [10], appropriately partitioned into two parts. The eigenvalues in T_1 form a *cluster* when

$$\|X_1 T_1(\alpha, \beta)^{-1} Z_1\| \gg \|X_2 T_2(\alpha, \beta)^{-1} Z_2\| \leq \epsilon \|X_1 T_1(\alpha, \beta)^{-1} Z_1\| \quad (4)$$

for some small constant ϵ. Consequently, we have

$$\|L(\alpha, \beta)^{-1}\| \leq (1 + \epsilon) \|X_1 T_1(\alpha, \beta)^{-1} Z_1\| \ .$$

Similar arguments and techniques to those in proving Theorem 2.1 can then be applied to $L(\alpha, \beta) + \delta L(\alpha, \beta)$ so that

$$(1 + \epsilon) \|X_1 T_1(\alpha, \beta)^{-1} Z_1\| \, \|\delta L(\alpha, \beta)\| \geq \|L(\alpha, \beta)^{-1} \delta L(\alpha, \beta)\| \geq 1$$

and

$$(1 + \epsilon) \kappa_1 \|\delta L(\alpha, \beta)\| \geq \|T_1(\alpha, \beta)^{-1}\|^{-1} \ , \quad \kappa_1 \equiv \|X_1\| \, \|Z_1\| \ .$$

Replacing $\|T_1(\alpha, \beta)^{-1}\|$ by an upper bound (as in the Appendix when T_1 is in Jordan or Kronecker form) yields similar results to those in Theorem 2.1, for the cluster in T_1 rather than the whole spectrum $\sigma(L)$. Here p will be the size of the largest Jordan block associated with the cluster in T_1. Ignoring higher order terms in ϵ, the perturbation results will then involve κ_1 instead of κ. The price to pay for the sharper result is the restriction that the perturbation δL has to be small (in the sense of (4)), contrary to its arbitrariness in Theorem 2.1.

(d) In (c) above, when T_1 contains a simple eigenvalue, κ_1 will be the product of the norms of the corresponding left- and right-eigenvecvtors. Similarly for a group of multiple eigenvalues, the corresponding condition number will be the product of the norms of the corresponding left- and right-eigenvectors (or deflating subspaces). Similar condition numbers can be obtained for clusters of eigenvalues.

(e) Obviously for large perturbations with $\theta_1 > 1$, $\max\{\theta_1, \theta_1^{1/p}\} = \theta_1$. Similarly when $\theta_1 < 1$, which is usually the case in (c) above, the maximum occurs at $\theta_1^{1/p}$. Furthermore, when the perturbation is asymptotically small, p in (2) equals the size of the Jordan block associated with (α_k, β_k) where the minimum in $s_{(\alpha,\beta)} \equiv \min_i \{|\alpha\beta_i - \beta\alpha_i|\}$ occurs at $i = k$. (In fact, a perturbation can be considered "small" when this correct pairing occurs; see the proof in Theorem 5.2.) Notice that the pth root is a common feature in perturbation results for multiple eigenvalues.

(f) A feature of the Bauer-Fike type perturbation result is that one starts with a perturbed eigenvalue (α, β) and its spectral variation from a nearby unperturbed eigenvalue (α_i, β_i) is bounded. As the size of the perturbation is unrestricted, there may well be unperturbed eigenvalues not paired up with any perturbed eigenvalues.

3 Bauer-Fike Theorem for Palindromic Pencils

For the pencil $\lambda Z - Z^\star$, we can work from the Kronecker canonical form

$$Q_1^\star (\lambda Z - Z^\star) Q_2 = \lambda \Lambda_+ - \Lambda_- \ , \quad Q_1 = [P_1, \ P_2] \ , \quad Q_2 = [P_2, \ P_1]$$

where

$$\Lambda_+ = \begin{bmatrix} \Lambda & \\ & I \end{bmatrix} \ , \quad \Lambda_- = \begin{bmatrix} I & \\ & \Lambda \end{bmatrix} \ , \quad \Lambda = \mathrm{diag}\{J_1, \cdots, J_N\}$$

with J_i being the Jordan block for λ_i on or inside the unit circle.

We have the following Bauer-Fike perturbation result.

Theorem 3.1 *Consider the palindromic pencil $L \equiv \beta Z - \alpha Z^\star$ with the above Kronecker canonical form. Let $\tilde{Z} = Z + \delta Z$, $\tilde{L} \equiv \beta \tilde{Z} - \alpha \tilde{Z}^\star$, $(\alpha_i, \beta_i) \in \sigma(L)$ and $(\alpha, \beta) \in \sigma(\tilde{L})$, with the scaling $|\alpha_i|^2 + |\beta_i|^2 = 1 = |\alpha|^2 + |\beta|^2$. The spectral variation of $\widetilde{L}$ from L is defined as*

$$s_L(\widetilde{L}) \equiv \max_{(\alpha,\beta)} \{s_{(\alpha,\beta)}\} \ , \quad s_{(\alpha,\beta)} \equiv \min_i \{|\alpha\beta_i - \beta\alpha_i|\} \ .$$

Then for any Holder norm $\|\cdot\|$, we have

$$s_{(\alpha,\beta)} \leq \max\{\theta_2, \theta_2^{1/p}\} \ , \quad \theta_2 \equiv c_2 \kappa_2 \ \|\delta Z\|$$

with κ_2 *as defined in (6),* p *being the size of the largest Jordan block in* Λ *and* $c_2 = 2\sqrt{2(|\alpha|^2 + |\beta|^2)} = 2\sqrt{2}$*.*

Also, we have

$$s_L(\widetilde{L}) \leq \max\{\theta_2, \theta_2^{1/p}\} \ .$$

Proof. Applying the techniques in [5], we consider the singular matrix

$$\begin{aligned} & Q_1^\star \left[\beta(Z + \delta Z) - \alpha(Z + \delta Z)^\star\right] Q_2 \\ &= (\beta\Lambda_+ - \alpha\Lambda_-) \left[I + (\beta\Lambda_+ - \alpha\Lambda_-)^{-1} Q_1^\star(\beta\delta Z - \alpha\delta Z^\star)Q_2\right] \end{aligned}$$

which implies

$$\kappa_2 \|(\beta\Lambda_+ - \alpha\Lambda_-)^{-1}\| \ \|\beta\delta Z - \alpha\delta Z^\star\| \geq 1 \tag{5}$$

where

$$\kappa_2 \equiv \|Q_1\| \ \|Q_2\| \ . \tag{6}$$

From Apendix I, we have the upper bound

$$\|(\beta\Lambda_+ - \alpha\Lambda_-)^{-1}\| \leq c_0 \max\{|z_i|^{-1}, |z_i|^{-p_i}\}$$

where $z_i \equiv \alpha\beta_i - \beta\alpha_i$, p_i is the size of the Jordan block associated with (α_i, β_i) and $c_0 \leq 2$. Substitute this bound into (5), we obtain

$$\kappa_2 c_2 \ \|\delta Z\| \ \max\{|z_i|^{-1}, |z_i|^{-p_i}\} \geq 1 \quad \Rightarrow \quad \min\{|z_i|, |z_i|^{p_i}\} \leq \kappa_2 c_2 \|\delta Z\| \ .$$

The results in the Theorem then follow.

■

Comments:

(a) Note that for the 2-norm or the F-norm (Frobenius), we have

$$\theta_2 \leq \sqrt{2(|\alpha|^2 + |\beta|^2)} \ p \ \|\delta Z\| \ .$$

With the scaling $|\alpha|^2 + |\beta|^2 = 1 = |\alpha_i|^2 + |\beta_i|^2$, $s_{(\alpha,\beta)}$ becomes the chordal metric [5, 11] and $\theta_2 \leq \sqrt{2} \ p \ \|\delta Z\|$. Although the F-norm is not a Holder norm, the corresponding results can be deduced from the 2-norm results.

(b) Comments similar to those in (c)–(f) after Theorem 2.1 apply for Thorem 3.1. The proof for the results for clusters of eigenvalues is similar, using the partitioning of an appropriate resolvent, and will not be repeated here.

(c) Details associated with palindromic "linearizations" can be found in [16]. Obviously, results in this Section are applicable to general palindromic pencils which may not be linearizations of matrix polynomials.

4 Bauer-Fike Theorem for Anti-Triangular Form

From [18], we have the following anti-triangular canonical form for $\star = T$:

Theorem 4.1 *Let $Z - \lambda Z^\star$ be a regular $n \times n$ palindromic pencil. There exists a unitary $U \in \mathcal{C}^{n\times n}$ such that $U^\star Z U = (m_{ij})$ with $m_{ij} = 0$ $(i + j \leq n + 1)$ (i.e., $U^\star Z U$ is anti-triangular, with zero elements in the upper left corner).*

Note that the result for $\star = H$ can easily be obtained by extending the proofs of Lemma 2.2 and Theorem 2.3 in [18]. Note also that we are only interested in the case when $Z - \lambda Z^\star$ is regular, which is not the case in [18].

The eigenvalues of the palindromic pencil $Z - \lambda Z^\star$ are:

$$\frac{m_{1n}}{m_{n1}^\star}, \frac{m_{2,n-1}}{m_{n-1,2}^\star}, \cdots, \frac{m_{i,n-i+1}}{m_{n-i+1,i}^\star}, \cdots, \frac{m_{n-i+1,i}}{m_{i,n-i+1}^\star}, \cdots, \frac{m_{n-1,2}}{m_{2,n-1}^\star}, \frac{m_{n1}}{m_{1n}^\star} .$$

Note that n will be even when considering a linearization of a palindromic quadratic pencil [16], but the results in this Section hold for any n.

Let N be the strict lower right triangular part of $U^\star Z U$. Reorganize the anti-triangular in Theorem 4.1 in upper triangular form:

$$\mathcal{P}_n U^\star (Z - \lambda Z^\star) U = (D_1 + N_1) - \lambda (D_2 + N_2) \tag{7}$$

with the order-reversing permutation matrix $\mathcal{P}_n = [e_n, e_{n-1}, \cdots, e_1]$, we have

$$\begin{aligned} D_2 &= \operatorname{diag}\{m_{1n}, m_{2,n-1}, \cdots, m_{n-1,2}, m_{n1}\} \\ D_1 = \mathcal{P}_n D_2 \mathcal{P}_n &= \operatorname{diag}\{m_{n1}, m_{n-1,2}, \cdots, m_{2,n-1}, m_{1n}\} \end{aligned}$$

where $N_1 = \mathcal{P}_n N$ and $N_2 = \mathcal{P}_n N^\star$ are strictly upper triangular.

Using the Schur-like form in (7), we can prove the following perturbation result for a palindromic pencil.

Theorem 4.2 *Consider the palindromic pencil $L \equiv \beta Z - \alpha Z^\star$ with N being the strictly lower right triangular part of its anti-triangular canonical form. Let $\tilde{Z} = Z + \delta Z$, $\tilde{L} \equiv \beta \tilde{Z} - \alpha \tilde{Z}^\star$, $(\alpha_i, \beta_i) \in \sigma(L)$ and $(\alpha, \beta) \in \sigma(\tilde{L})$. Assume the scaling $|\alpha|^2 + |\beta|^2 = 1 = |\alpha_i|^2 + |\beta_i|^2$. Then for any Holder norm $\|\cdot\|$, we have*

$$s_{(\alpha,\beta)} \leq \sqrt{2}\, c_0 \;\max\{\theta_3,\; c_3 \theta_3^{1/p}\} , \quad \theta_3 \equiv \|\delta Z\|$$

for some $p \leq n$ and $c_0 \equiv \min\{2,\ p\}$, $c_3 \equiv \sqrt{2}\,\|N\|^{1-1/p}$.

Also, we have

$$s_L(\widetilde{L}) \leq \sqrt{2}\, c_0 \max\{\theta_3,\; c_3 \theta_3^{1/p}\} .$$

Proof. Consider the singular matrix

$$\beta(Z+\delta Z)-\alpha(Z+\delta Z)^\star$$
$$=(U^\star)^{-1}\mathcal{P}_n\left[\beta(D_1+N_1+\mathcal{P}_nU^\star\delta ZU)-\alpha(D_2+N_2+\mathcal{P}_nU^\star\delta Z^\star U)\right]U^H$$
$$=(U^\star)^{-1}\mathcal{P}_n\left[\beta(D_1+N_1)-\alpha(D_2+N_2)\right]\cdot$$
$$\left\{I+\left[\beta(D_1+N_1)-\alpha(D_2+N_2)\right]^{-1}\mathcal{P}_nU^\star(\beta\delta Z-\alpha\delta Z^\star)U\right\}U^H\ .$$

It is easy to see that

$$\left\|\left[\beta(D_1+N_1)-\alpha(D_2+N_2)\right]^{-1}\right\|\ \|\beta\delta Z-\alpha\delta Z^\star\|\geq$$
$$\left\|\left[\beta(D_1+N_1)-\alpha(D_2+N_2)\right]^{-1}\mathcal{P}_nU^\star(\beta\delta Z-\alpha\delta Z^\star)U\right\|\geq 1\ . \qquad (8)$$

Note that $[\beta(D_1+N_1)-\alpha(D_2+N_2)]$ is assumed to be nonsingular, otherwise the results in the theorem become trivial.

With $z\equiv\min_{(\alpha_i,\beta_i)}\|\beta D_1-\alpha D_2\|=\min_{(\alpha_i,\beta_i)}|\beta\alpha_i-\alpha\beta_i|$, $\tilde{D}\equiv\beta D_1-\alpha D_2$ and $\tilde{N}\equiv\beta N_1-\alpha N_2$, we have

$$\begin{aligned}M &\equiv \left[\beta(D_1+N_1)-\alpha(D_2+N_2)\right]^{-1}\\ &= \left[I-(\beta D_1-\alpha D_2)^{-1}(\beta N_1-\alpha N_2)\right]^{-1}(\beta D_1-\alpha D_2)^{-1}\\ &= (I-\tilde{D}^{-1}\tilde{N})^{-1}\tilde{D}^{-1}\ .\end{aligned}$$

As $\tilde{D}^{-1}\tilde{N}$ is nilpotent, there exists some $p\leq n$ such that

$$(I-\tilde{D}^{-1}\tilde{N})^{-1}=I+\tilde{D}^{-1}\tilde{N}+\cdots+(\tilde{D}^{-1}\tilde{N})^{p-1}$$

and we obtain

$$\|M\|\leq\|\tilde{N}\|^{-1}\eta^{-1}\equiv z^{-1}(1+\|\tilde{N}\|z^{-1}+\cdots+\|\tilde{N}\|^{p-1}z^{-p+1})\ . \qquad (9)$$

With $x\equiv\|\tilde{N}\|^{-1}z$, we have the polynomial $P(x)\equiv x^p-\eta(1+x+\cdots+x^{p-1})=0$, as in the Appendix or [5]. The only positive root x of P satisfies

$$x\equiv\|\tilde{N}\|^{-1}z\leq c_0\ \max\{\eta,\ \eta^{1/p}\}\ .$$

As (8) and (9) imply

$$\sqrt{2}\ \|\delta Z\|\geq\|\beta\delta Z-\alpha\delta Z^\star\|\geq\|M\|^{-1}\geq\|\tilde{N}\|\eta$$

this and the upper bound in (9) then lead to

$$z\leq c_0\|\tilde{N}\|\ \max\{\eta,\eta^{1/p}\}\leq\sqrt{2}\,c_0\ \max\{\|\delta Z\|,\ \|\tilde{N}\|^{1-1/p}\ \|\delta Z\|^{1/p}\}$$

As $\|\tilde{N}\|\leq\sqrt{2}\ \|N\|$, the results in the Theorem then follow.

■

With the chosen scaling $|\alpha|^2+|\beta|^2=1=|\alpha_i|^2+|\beta_i|^2$, $s_{(\alpha,\beta)}$ equals the chordal metric. Note also that p is the integer for which $\tilde{N}^k\neq 0$ $(0\leq k<p)$ and $\tilde{N}^p=0$.

5 Bauer-Fike Theorem for Semi-Schur Anti-Triangular Form

A refinement of Theorem 4.2, with a smaller value for p, can be proved. We first refine the decomposition in Theorem 4.1.

Theorem 5.1 *Let $Z - \lambda Z^\star$ be a regular palindromic pencil. There exist nonsingular U, $V \in \mathcal{C}^{n\times n}$ such that*

$$V^\star Z U = \textit{anti-diag}\{M_1, \cdots, M_r\}\ , \quad M_j = D_j + N_j \tag{10}$$

where M_j is anti-triangular with anti-diagonal elements in $D_j = \textit{anti-diag}\{\lambda_j, \cdots, \lambda_j\}$.

Proof. The proof is similar to the standard transformation of a Schur decomposition to the corresponding Jordan canonical form. It is sufficient to show that it is possible to transform the anti-triangular form $(\tilde{U}^\star Z\tilde{U},\ \tilde{U}^\star Z^\star \tilde{U})$ in Theorem 4.1 to anti-block diagonal form, so that

$$\begin{bmatrix} I & 0 \\ P^\star & I \end{bmatrix} \tilde{U}^\star Z \tilde{U} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ P^\star & I \end{bmatrix} \begin{bmatrix} 0 & T_1 \\ T_2 & T_{12} \end{bmatrix} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} = \begin{bmatrix} & T_1 \\ T_2 & \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ P^\star & I \end{bmatrix} \tilde{U}^\star Z^\star \tilde{U} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ P^\star & I \end{bmatrix} \begin{bmatrix} 0 & T_2^\star \\ T_1^\star & T_{12}^\star \end{bmatrix} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} = \begin{bmatrix} & T_2^\star \\ T_1^\star & \end{bmatrix}$$

when T_1 and T_2 have nonintersecting spectra. Multiplying out the above equations produces

$$\phi(P, Q) \equiv (T_2 Q + P^\star T_1,\ T_1^\star Q + P^\star T_2^\star) = -(T_{12},\ T_{12}^\star) \tag{11}$$

which is uniquely solvable [3].

■

Similar to Theorem 5.1 but with p being the bounded by the maximum size of M_j, we now have the following refined version of Theorem 4.2:

Theorem 5.2 *Consider the palindromic pencil $L \equiv \beta Z - \alpha Z^\star$ with its semi-Schur anti-triangular canonical form (10). Let $\tilde{Z} = Z + \delta Z$, $\tilde{L} \equiv \beta\tilde{Z} - \alpha\tilde{Z}^\star$, $(\alpha_i, \beta_i) \in \sigma(L)$ and $(\alpha, \beta) \in \sigma(\tilde{L})$. Assume the scaling $|\alpha|^2 + |\beta|^2 = 1 = |\alpha_i|^2 + |\beta_i|^2$. Then for any Holder norm $\|\cdot\|$, we have*

$$s_{(\alpha,\beta)} \le \sqrt{2}\, c_0\ \max\{\theta_4,\ c_4\theta_4^{1/p}\}\ , \quad \theta_4 \equiv \kappa_4 \|\delta Z\|$$

for $\kappa_4 \equiv \|U\|\,\|V\|$, $c_0 \equiv \min\{2,\ p\}$ *and* $c_4 \equiv \sqrt{2}\ \max_j \|N_j\|^{1-1/p}$.

For sufficiently small perturbations, we have p *being the size of the Schur block* M_j *associated with* $(\alpha_i,\ \beta_i)$*. In general, we have* $p = p^*$*, the maximum size of the Schur blocks* M_j*.*

We also have

$$s_L(\widetilde{L}) \leq \sqrt{2}\, c_0\ \max\{\theta_4,\ c_4 \theta_4^{1/p^*}\}\ .$$

Proof. The proof is exactly the same as that for Theorem 4, except $\|M\|$ in (9) equals to the maximum of the norms of its diagonal blocks. The same argument then follows in a similar fashion as in the Proof of Theorem 4, using the diagonal block at which the maximum occurs.

When the perturbation is small enough, this maximum (nearly infinite) occurs at the same block associated with (α_k, β_k) at which $\min_i\{|\alpha\beta_i - \beta\alpha_i|\}$ occurs. This gives a sharper perturbation result, with a smaller p being the size of the diagonal block associated with (α_k, β_k).

■

Comments similar to those in (c)–(f) after Theorem 2.1 apply for the above Theorem. Similar to Theorem 3.1, when we consider sufficiently small perturbations, there will be a 1-1 correspondence between the original and perturbed eigenvalues. The above perturbation bounds can be proved for a particular eigenvalue (α_i, β_i) with the condition number κ_4 replacable by $\|U_j\|\,\|V_j\|$. In addition, we can consider a group of neighbouring eigenvalues together, instead of one particular eigenvalue. This will increase p, or the size of the corresponding semi-Shcur block M_j, but will improve the condition of the linear operator ϕ in (11) as well as κ_4.

6 Perturbation by Differentiation

The results in this Section are quoted from [7].

Without establishing differentiability or the existence of asymptotic expansions (which can be achieved using the implicit function approach), perturbation results can be obtained by simple differentiation. See [1] for more details on this approach.

For some fixed $z \neq 0$, consider the palindromic eigenvalue problem

$$P(\lambda, \rho)x(\rho) = 0\ , \quad P(\lambda, \rho) \equiv \lambda(\rho)^2 A_1^\star(\rho) + \lambda(\rho)A_0(\rho) + A_1(\rho)$$

with the scaling $z^\star x(\rho) - 1 = 0$, where ρ is the perturbation parameter, $A_0(0) = A_0$ and $A_1(0) = A_1$. We shall use the subscripts $(\cdot)_\rho$ and $(\cdot)_\lambda$

to denote the corresponding partial derivatives. For a simple eigenvalue λ, differentiation produces, at $\rho = 0$:

$$\lambda_\rho = -\frac{y^\star P_\rho x}{y^\star P_\lambda x} = -\frac{y^\star P_\rho x}{y^\star (2\lambda A_1^\star + A_0) x} \tag{12}$$

and

$$P x_\rho = -(\lambda_\rho P_\lambda + P_\rho) x \ , \quad z^\star x_\rho = 0 \ .$$

Choosing $z = y(0)$ (the left-eigenvector corresponding to $\lambda(0)$), we have, at $\rho = 0$:

$$x_\rho = -P^\dagger (\lambda_\rho P_\lambda + P_\rho) x$$

where $P^\dagger$ denotes the Penrose generalized inverse [11].

The usual conclusions can be drawn — the right-eigenvector x will be rotated through a big angle, even for a small perturbation, when $\|P^\dagger\|$ is large, i.e., when the separation between λ and other eigenvalues is fine. This happens, of course, when the assumption of simplicity for the eigenvalue is near collapsing.

Note that for palindromic eigenvalues problems with $\star = T$, $\lambda = \pm 1$ are multiple and non-differentiable, and a more sophisticated approach, like the one in [4], is required.

For perturbation results obtained through the application of Sun and Stewart's approach [9, 20] in terms of the implicit function theorem, see [7]. Asymptotic perturbation series for the eigenvalues and the deflating subspaces have been obtained.

7 Conclusions

Bauer-Fike type perturbation results for general matrix polynomials, palindromic linearizations and the (semi-Schur) anti-triangular canonical forms are discussed, for perturbations of arbitrary size. These perturbation results complements those for asymptotic perturbations in [7]. Consistent results for simple eigenvalues and the corresponding eigenvetors are obtained using simple differentiation. These results indicate, not surprisingly, that the perturbations of an eigenvalue λ and its corresponding deflating subspace $\mathcal{S}_\lambda$, respectively, are proportional to the size of the perturbation and the reciprocal of the gap between $\mathcal{S}_\lambda$ and other deflating subspaces. Condition numbers are typically proportional the products of the norms of the left- and right-eigenvectors or deflating subspaces.

Appendix I: Bounding $\|(\beta\Lambda_+ - \alpha\Lambda_-)^{-1}\|$

For any Holder norm, with

$$\Lambda_+ = \begin{bmatrix} \Lambda & \\ & I \end{bmatrix} , \quad \Lambda_- = \begin{bmatrix} I & \\ & \Lambda \end{bmatrix} , \quad \Lambda = \operatorname{diag}\{J_1, \cdots, J_N\}$$

and the eigenvalues of J_i all on or inside the unit circle, we have $\|(\beta\Lambda_+ - \alpha\Lambda_-)^{-1}\| = \max_i \|M_i\|$, where $M_i \equiv (\alpha I - \beta J_i)^{-1} \in \mathcal{R}^{p_i \times p_i}$. Thus it is sufficient to consider the bound for $\|M_i\|$. Let $z_i \equiv \alpha\beta_i - \beta\alpha_i$ with α_i and β_i being diagonal elements of I and J_i, respectively. (Here, $\alpha_i = 1$ and $|\beta_i| \leq 1$; the case when $M_i \equiv (\alpha J_i - \beta I)^{-1}$ can be treated similarly, and the symmetric notations used here can then be adapted easily.) We have the Toeplitz matrix

$$M_i \equiv \begin{bmatrix} z_i & -\beta & & & \\ & z_i & -\beta & & 0 \\ & & \ddots & \ddots & \\ & 0 & & z_i & -\beta \\ & & & & z \end{bmatrix}^{-1} = \begin{bmatrix} z_i^{-1} & \beta z_i^{-2} & \beta^2 z_i^{-3} & \cdots & \beta^{p_i-1} z_i^{-p_i} \\ & z_i^{-1} & \beta z_i^{-2} & \cdots & \beta^{p_i-2} z_i^{-p_i+1} \\ & & z_i^{-1} & \ddots & \vdots \\ & 0 & & \ddots & \beta z_i^{-2} \\ & & & & z_i^{-1} \end{bmatrix} .$$

We then have

$$\|M_i\| \leq c_1 \cdot \max\{|z_i|^{-1},\ |z_i|^{-p_i}\} .$$

Finally, p_i and $|z_i|$ can respectively be replaced by $p \equiv \max p_i$ and $z \equiv \min_i |\alpha\beta_i - \beta\alpha_i|$, with $c_1 = p$.

Alternatively, potentially sharper bounds with can be obtained, with c_1 replaced by c_0. With $M_i^{-1} = z_i \left[I_{q_i} - z_i^{-1} N^{(i)}\right]$ for a nilpotent $N^{(i)}$ such that $\left[N^{(i)}\right]^{q_i} = 0$, we have

$$M_i = z_i^{-1} \sum_{j=0}^{q_i-1} z_i^{-1} \left[N^{(i)}\right]^j , \quad \|M_i\| \leq \eta^{-1} \equiv |z_i|^{-1} \sum_{j=0}^{q_i-1} |z_i|^{-j} .$$

For simplicity, let $x = |z_i|$ and $m = q_i$, the above definition of η leads to the polynomial

$$P_m(x) \equiv x^m - \eta(1 + x + \cdots + x^{m-1}) .$$

Descartes' sign rule (*La Géométrie* 1637) then implies that $P_m(x)$ has at most one positive real root. As $P_m(0) = -\eta < 0$ and $P_m(x) > 0$ as $x \to \infty$, any positive number x^* for which $P_m(x^*) > 0$ is an upper bound of the unique real positive root of $P_m(x)$. Simple inspection leads to the upper bounds $x^* = c_0\eta$ when $\eta > 1$ and $x^* = c_0\eta^{1/m}$ when $\eta \leq 1$, with $c_0 = \min\{2,\ m\}$. Consequently, $c_0 = 1$ when $m = 1$ (with $x = \eta$) and $c_0 = 2$ when $m > 1$.

The details are as follows. When $c_0\eta > \eta \geq 1$ and $m > 1$, we have

$$\begin{aligned} P_m(c_0\eta) &= (c_0\eta)^m - \eta\frac{(c_0\eta)^m - 1}{c_0\eta - 1} \\ &= \frac{c_0^{m+1}\eta^{m+1} - c_0^m\eta^m - c_0^m\eta^{m+1} + \eta}{c_0\eta - 1} \\ &= \frac{(c_0^m\eta^{m+1} - c_0^m\eta^m) + (c_0^{m+1}\eta^{m+1} - c_0^m\eta^{m+1} - c_0^m\eta^{m+1}) + \eta}{c_0\eta - 1} \geq 0 \end{aligned}$$

as $c_0^{m+1}\eta^{m+1} - c_0^m\eta^{m+1} - c_0^m\eta^{m+1} = (c_0 - 2)c_0^m\eta^{m+1} = 0$. Thus $c_0\eta$ is an upper bound of the root x of P_m when $\eta \geq 1$.

When $\eta < 1$ and $m > 1$, we have

$$\begin{aligned} P_m(c_0\eta^{1/m}) &\geq c_0^m\eta - \eta(1 + c_0 + \cdots + c_0^m) \\ &= c_0^m\eta - \eta\frac{c_0^m - 1}{c_0 - 1} = \eta > 0 \ . \end{aligned}$$

Thus $c_0\eta^{1/m}$ is an upper bound of the root x of P_m when $\eta < 1$.

Finally, we have

$$|z_i| \leq c_0 \max\{\eta, \eta^{-q_i}\} \quad \Rightarrow \quad \|M_i\| \leq \eta^{-1} \leq c_0 \ \max\{|z_i|^{-1}, |z_i|^{-q_i}\}$$

and the result follows, with c_1 replaced by the sharper c_0.

References

[1] A.L. ANDREW, K.-W.E. CHU AND P. LANCASTER, Derivatives of eigenvalues and eigenvectors of nonlinear eigenvalue problems, *SIAM J. Matrix Anal. Applic.*, **14** (1993), pp. 903–926.

[2] S. BORA, Structured eigenvalue condition number and backward error of a class of polynomial eigenvalue problems, Technical Report **417**, DFG Research Centre Matheon, Berlin, 2007.

[3] K.-W.E. CHU, The solution of the matrix equations $AXB - CXD = E$ and $(YA - DZ, YC - BZ) = (E, F)$, *Lin. Alg. Applic.*, **93** (1987), pp. 93–105.

[4] K.-W.E. CHU, On multiple eigenvalues of matrices depending on several parameters, *SIAM J. Numer. Anal.*, **27** (1990), pp. 1368–1385.

[5] E.K.-W. CHU, Perturbation of eigenvalues for matrix polynomials via the Bauer-Fike theorems, *SIAM J. Matrix Anal. Applic.*, **25** (2003), pp. 551–573.

[6] E. K.-W. CHU, T.-M. HWANG, W.-W. LIN, AND C.-T. WU, Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms, *J. Comput. Appl. Maths.*, **219** (2007), pp. 237–252. (doi: 10.1016/j.cam.2007.07.016).

[7] E. K.-W. CHU, W.-W. LIN, AND C.-S. WANG, Asymptotic perturbation of palindromic eigenvalue problems, *Preprints and Reprints*, National Centre of Theoretical Sciences, National Tsing Hua University, Taiwan, August 2007. (To appear in *Taiwanese J. Maths.*)

[8] J.-P. DEDIEU AND F. TISSEUR, Perturbation theory for polynomial eigenvalue problems in homogeneous form, *Lin. Alg. Applic.*, **358** (2003), pp. 71–94.

[9] L. ELSNER AND J.G. SUN, Perturbation theorems for the generalized eigenvalue problem, *Lin. Alg. Applic.*, **48** (1982), pp. 341–357.

[10] I. GOHBERG, P. LANCASTER AND L. RODMAN, *Matrix Polynomials*, Academic Press, New York, 1982.

[11] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, 3rd Ed., Johns Hopkins University Press, Baltimore, 1996.

[12] A. HILLIGES, *Numerische Lösung von quadratischen Eigenwertproblemen mit Anwendungen in der Schienendynamik*, Master Thesis, Technical University Berlin, Germany, July 2004.

[13] A. HILLIGES, C. MEHL AND V. MEHRMANN, On the solution of palindromic eigenvalue problems, *Proceedings 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*, Jyväskylä, Finland, 2004.

[14] C.F. IPSEN, Accurate eigenvalues for fast trains, *SIAM News*, **37** (2004).

[15] D. KRESSNER, M.J. PELEZ AND J. MORO, Structured Hölder condition numbers for multiple eigenvalues, UMINF report, Department of Computer Sciences, Ume University, Sweden, 2006. (Revised January 2008; available at `http://www.sam.math.ethz.ch/ kressner/pubs.php`)

[16] D.S. MACKEY, N.MACKEY, C. MEHL AND V. MEHRMANN, Linearization spaces for matrix polynomials, *SIAM J. Matrix Applic.* , **28** (2006), pp. 971–1004.

[17] D.S. MACKEY, N. MACKEY, C. MEHL AND V. MEHRMANN, Structured polynomial eigenvalue problems: good vibrations from good linearizations, *SIAM J. Matrix Applic.*, **28** (2006), pp. 1029–1051.

[18] D.S. MACKEY, N. MACKEY, C. MEHL AND V. MEHRMANN, Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form, Technical Report **409**, DFG Research Centre Matheon, Berlin, 2007. (To appear in *Numer. Lin. Alg. Appl.*)

[19] MATHWORKS, *MATLAB User's Guide*, 2002.

[20] G.W. STEWART AND J.G. SUN, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[21] F. TISSEUR, Backward error and condition of polynomial eigenvalue problems, *Lin. Alg. Applic.*, **309** (2000), pp. 339–361.

[22] F. TISSEUR AND K. MEERBERGEN, A survey of the quadratic eigenvalue problem, *SIAM Rev.*, **43** (2001), pp. 234–286.