BACKWARD PERTURBATION AND SENSITIVITY ANALYSIS OF STRUCTURED POLYNOMIAL EIGENVALUE PROBLEM

by

Bibhas Adhikari

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI-781039, INDIA

December 2008

BACKWARD PERTURBATION AND SENSITIVITY ANALYSIS OF STRUCTURED POLYNOMIAL EIGENVALUE PROBLEM

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Bibhas Adhikari

(Roll Number: 04612301)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

December 2008

TH-789_04612301

CERTIFICATE

It is certified that the work contained in the thesis titled **"Backward Perturbation and Sensitivity Analysis of Structured Polynomial Eigenvalue Problem"** by **Bibhas Adhikari**, a student in the Department of Mathematics, Indian Institute of Technology Guwahati for the award of the degree of Doctor of Philosophy has been carried out under my supervision and this work has not been submitted elsewhere for a degree.

Prof. Rafikul Alam
Department of Mathematics
Indian Institute of Technology Guwahati

TH-789_04612301

Dedicated to

Davis-Kahan-Weinberger dilation Theorem

TH-789_04612301

Abstract

The main theme of the thesis is structured perturbation and sensitivity analysis of structured polynomial eigenvalue problem.

Structured mapping problem naturally arises when analyzing structured backward perturbation of structured eigenvalue problem. Given two matrices X and B of same size, the structured mapping problem requires to find a "structured" matrix A, if any, having the smallest norm such that $AX = B$. We provide a complete solution of structured mapping problem. More generally, we provide a complete solution of the structured inverse least-squared problem (SILSP):

$$\min_A \|AX - B\|_F,$$

where the minimum is taken over "structured" matrices. As a consequence of structured mapping problem, we determine structured backward errors of approximate invariant subspaces of structured matrices. We also analyze structured pseudospectra of structured matrices.

Next, we undertake a detailed structured backward perturbation analysis of structured matrix polynomials and derive explicit computable expressions for structured backward errors of approximate eigenelements. We analyze structured pseudospectra of structured matrix polynomials and establish a partial equality between unstructured and structured pseudospectra, which plays an important role in solving certain distance problems associated with structured polynomials. We also derive relatively simple expressions for structured condition numbers of simple eigenvalues of structured matrix polynomials, which play an important role in analyzing sensitivity of eigenvalues of structured polynomial eigenvalue problem.

Generally, a polynomial eigenvalue problem is "linearized" first and then solved by a backward stable algorithm. However, the eigenvalues of the resulting linear problem is usually more sensitive to perturbation than the original problem. Moreover, a polynomial admits infinitely many linearizations. The same holds true for structured polynomials as well. Therefore, for computational purposes, it is of paramount importance to identify potential structured linearizations which are as well conditioned as possible. With the help of structured backward perturbation analysis and structured condition numbers of eigenvalues, we identify "good" structured linearizations which guarantee almost as accurate solutions as that of the original polynomial eigenvalue problem.

TH-789_04612301

না গো, এই যে ধূলা আমার না এ।
তোমার ধূলার ধরার 'পরে উড়িয়ে যাব সন্ধ্যাবায়ে॥
দিয়ে মাটি আগুন জ্বালি রচলে দেহ পূজার থালি—
শেষ আরতি সারা করে ভেঙে যাব তোমার পায়ে॥
ফুল যা ছিল পূজার তরে
যেতে পথে ডালি হতে অনেক যে তার গেছে পড়ে।
কত প্রদীপ এই থালাতে সাজিয়েছিলে আপন হাতে—
কত যে তার নিবল হাওয়ায়, পৌঁছলনা চরণ-ছায়ে॥

রবীন্দ্রনাথ ঠাকুর

TH-789_04612301

Acknowledgement

First of all I would like to express my extreme gratitude to Prof. Rafikul Alam for making the experience of working on this thesis worthwhile. Prof. Alam's expertise and clarity of thought in conducting research as well as in its presentation have strongly motivated me. Moreover, discussions with him have sparked off my ideas and helped me to discover my field of interest in pursuing research.

I would like to thank Indian Institute of Technology Guwahati (IITG) for the facilities provided to me during my research work and to the Ministry of Human and Resource Development, Govt. of India, for offering me financial assistance for the completion of my thesis work.

Part of this work was done when I was visiting Prof. Dr. Daniel Kressner, SAM, ETH Zurich. The financial support by the Indo Swiss Bilateral Research Initiative (ISBRI), which made the visit possible, is gratefully acknowledged. I am indebted to Dr. Daniel Kressner for his help and encouragement, which contributed greatly in the improvement of the chapter 6 of the thesis.

I owe my thanks to the doctoral committee members Dr. Rajen K. Sinha, Dr. Sukanta Pati and Dr. Shreemayee Bora for their valuable suggestions. I would also like to thank Prof. Volker Mehrmann of TU Berlin for the valuable discussions I had with him when he visited IITG. I am grateful to Prof. Yimin Wei of Fudan University, China for the valuable discussions I had with him during the workshop in CUHK, Honk Kong and also for providing with a lot of literature and research articles, which are not available in our institute. My sincere thanks goes to Mr. Peter Kauf, Mr. Prosenjit Bhowmik and Mr. Ashis Roy for helping me in various ways for making my stay at ETH Zurich enjoyable and comfortable.

There are many people who have stimulated my thoughts. In particular, I would like to thank Dr. K. V. Srikanth and Dr. Sriparna Bandopadhyay of the Department of Mathematics, IIT Guwahati for answering my countless queries.

I want to convey my sincere thanks to Dr. Shankar De of Nataji Nagar Day College Kolkata, Dr. Sachin Boksi of Belur Vidyamandir Howrah, Dr. Alokes Maity of Presidency College Kolkata for inspiring me to pursue my career in Mathematics. I am grateful to Prof. S. K. Acharya, Prof. S. Ganguly and Prof. M. N. Mukherjee of the Department of Pure Mathematics, University of Calcutta and Dr. Mahuya Datta of ISI Calcutta, for motivating me in pursuing research.

Thanks are also due to my friends and seniors Dr. Sunil K. Maity, Dr. Swapan K. Pandit and Dr. Safique Ahmad for a lot of help and encouragement. I offer my thanks to our lab assistant Mr. Shantanu Majumdar for technical support. Sridhar Samal, Manoj Boro of Department of Mathematics deserve special thanks for their assistance in all official matters.

I would also like to thank my friends I have at IIT Guwahati whose company provided academic help and a welcome distraction from academic works. Notably Bidyut, Avijit, Bitta, Deba, Laha, Kaushik, Jugal, Sunil, Achin, Tara, Barun, Tarakanta, Bhupen, Sasmita, Madhusmita, Martha, Milan, Rajendra, Prabhanjan, Rajesh, Smruti, Cosmika, Akhilesh, Raju, Rupam, Debajit, Shuvam, Pushpendra, Shub, Ravi, Murli as well as sundry others who have shared many lively moments with me. Mahavir, Mahanty, Prasanta and Sourav Adhikari deserve a special thank for being with me all the time.

I would like to express my regards to my mother, Mrs. Archana Adhikari, to whom I owe

TH-789_04612301

everything I have achieved and whatever I have become; my father, Mr. Provas Adhikari, for always having faith in me and inspiring me; and my sisters, Dolon and Ananya, who were always my silent supporters during this endeavour and beyond.

Last but not the least, to my dear Papri, who came along with me in every step of this journey and was always there for me whenever I needed her most.

December 2008 **(Bibhas Adhikari)**

TH-789_04612301

Contents

List of Tables ix

1 Introduction **1**

1.1 Introduction 1

1.2 Preliminaries 3

1.2.1 Davis-Kahan-Weinberger dilation theorem 5

1.2.2 Matrix polynomials and linearizations 7

1.2.3 Structured matrices and matrix polynomials 11

2 Structured mapping problem **15**

2.1 Introduction 15

2.2 Solution of structured mapping problem 16

2.3 Structured backward errors of eigenelements 24

2.3.1 Structured pseudospectra 25

2.4 Structured backward error for approximate invariant subspaces of structured matrices 27

2.5 Structured inverse least square problem 32

3 Structured backward errors and pseudospectra of structured matrix pencils **38**

3.1 Introduction 38

3.2 Structured matrix pencils 40

3.3 Frobenius norm and structured backward errors 43

3.4 Spectral norm and structured backward errors 53

3.5 Structured pseudospectra of structured pencils 61

4 Structured backward errors and linearizations for structured matrix polynomials **65**

4.1 Introduction 65

4.2 Structured matrix polynomials 67

TH-789_04612301

4.3 Structured backward error 68
4.4 Structured backward error and structured linearizations 80
4.5 Structured pseudospectra of structured matrix polynomials 91

5 Backward errors and linearizations for palindromic matrix polynomials 94
5.1 Introduction 94
5.2 Eigen-symmetry of palindromic polynomials 95
5.3 Structured backward error of approximate eigenpair 97
5.4 Structured backward error and palindromic linearizations 119
5.5 Pseudospectra of palindromic polynomials 123

6 Structured eigenvalue condition numbers and linearizations for matrix polynomials 126
6.1 Introduction 126
6.2 Structured condition numbers for matrix polynomials 127
6.2.1 Structured first-order perturbation sets 130
6.2.2 T-symmetric matrix polynomials 132
6.2.3 T-even and T-odd matrix polynomials 133
6.2.4 T-palindromic and T-anti-palindromic matrix polynomials 135
6.2.5 H-Hermitian matrix polynomials 137
6.3 Condition numbers for linearizations 138
6.4 Structured condition numbers for linearizations 140
6.4.1 T-symmetric matrix polynomials 140
6.4.2 T-even and T-odd matrix polynomials 141
6.4.3 T-palindromic matrix polynomials 149
6.4.4 H-Hermitian matrix polynomials and related structures 154
6.5 Summary and conclusions 154
6.6 Appendix 155

References 158

Indian Institute of Technology Guwahati

List of Tables

1.1 The Jordan and Lie algebras corresponding to various scalar products. 12

1.2 Overview of structured matrix polynomials. Note that $* \in \{T, H\}$ may denote either the complex transpose $(* = T)$ or the Hermitian transpose $(* = H)$. . . . 14

1.3 Structured linearizations of structured polynomials. Here $* \in \{T, H\}$. 14

2.1 Necessary and sufficient conditions for structured mapping problem. 17

2.2 Necessary and sufficient condition 23

2.3 Spectral symmetry of structured matrices. 24

3.1 Eigensymmetry of structured pencils and related conditions. 41

3.2 Numerical computation for structured backward error for structured pencils. . 62

4.1 Eigen-symmetry of structured polynomials. 67

4.2 Table for the admissible ansatz vectors for structured polynomials. 83

4.3 The choice of structured linearizations, where $\delta_v = \|\Lambda_{m-1}\|_2/|\Lambda_{m-1}^* v|$, $* \in \{T, H\}$ and v is the ansatz vector. 90

5.1 Eigensymmetry of palindromic polynomials. 96

5.2 Table for the admissible ansatz vectors for palindromic polynomials. 120

5.3 Partial equality of structured backward errors 123

5.4 The choice of structured linearizations for palindromic polynomials, where $\delta_v = \|\Lambda_{m-1}\|_2/|\Lambda_{m-1}^* v|$, $* \in \{T, H\}$ and v is the ansatz vector. 125

6.1 Parameters defining the ellipse (6.15) 131

6.2 Conditions the ansatz vector v needs to satisfy in order to yield a structured linearization $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ for a structured polynomial P. 140

6.3 Conditions the ansatz vector v needs to satisfy to yield a structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ such that $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$, where $\Sigma \in \mathbb{R}^{m \times m}$ is defined as $\Sigma = \mathrm{diag}\{(-1)^{m-1}, (-1)^{m-2}, \ldots, (-1)^0\}$ and $* \in \{T, H\}$. 142

TH-789_04612301

6.4 Recipes for choosing the ansatz vector v for a T-even or T-odd linearization L_e or L_o of a T-even matrix polynomial of degree m. Note that e_1 and e_m denote the 1st and mth unit vector of length m, respectively. 146

6.5 Conditions the ansatz vector v needs to satisfy to yield a structured linearization such that $(R \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. 149

6.6 Recipes for choosing the ansatz vector v for a T-palindromic or T-anti-palindromic linearization L_p or L_a of a T-palindromic matrix polynomial of degree m. Note that $\alpha_m = 2^{1/(m-1)}$. 153

TH-789_04612301

Chapter 1

Introduction

1.1 Introduction

Polynomial eigenvalue problem occurs in many practical applications. Often, a polynomial eigenvalue problem that occurs in practice has some distinctive structure, as a result of which the eigenvalues inherit certain spectral symmetry (see, [22, 44, 66, 68, 82] and the references therein). Thus numerical methods for solution of a structured polynomial eigenvalue problem that preserves spectral symmetry in the computed eigenvalues is highly desirable and is a challenging task [9, 10, 18, 44, 46, 71, 74].

We consider matrix polynomial of the form $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$, where $A_j \in \mathbb{C}^{n\times n}, j = 0 : m$. Then the polynomial eigenvalue problem is concerned with finding $(\lambda, x, y) \in \mathbb{C} \times \mathbb{C}^n \times \mathbb{C}^n$ such that

$$\mathrm{P}(\lambda)x = 0 \text{ and } y^H \mathrm{P}(\lambda) = 0, \tag{1.1}$$

where x and y are nonzero vectors known as right and left eigenvectors of P corresponding to the eigenvalue λ, respectively. However, due to the lack of a genuine polynomial eigensolver, the standard way of solving a polynomial eigenvalue problem of degree m is to solve an equivalent generalized eigenvalue problem of larger size. To be specific, an $n \times n$ polynomial P of degree m is converted into an "equivalent" linear polynomial

$$\mathrm{L}(\lambda) = \lambda X + Y, \;\; X, Y \in \mathbb{C}^{mn\times mn}$$

and numerically backward stable algorithm is employed to compute the eigenelements of L. By "equivalent" we mean to convert the polynomial P into a linear polynomial L satisfying

$$E(\lambda)\mathrm{L}(\lambda)F(\lambda) = \begin{bmatrix} \mathrm{P}(\lambda) & 0 \\ 0 & I_{(m-1)n} \end{bmatrix}, \text{ for all } \lambda \in \mathbb{C}$$

where $E(\lambda)$ and $F(\lambda)$ are unimodular polynomials. The linear polynomial L is called a linearization of P.

The most commonly used linearizations of a polynomial $\mathrm{P} = \sum_{j=0}^{m} z^j A_j$ are the block-companion forms, the first companion form $C_1(z) = zX_1 + Y_1$ and the second companion

TH-789_04612301

form $C_2(z) = zX_2 + Y_2$, where $X_1 = X_2 = \text{diag}(A_m, I_n, \ldots, I_n)$ and

$$Y_1 = \begin{bmatrix} A_{m-1} & A_{m-2} & \ldots & A_0 \\ -I_n & 0 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & -I_n & 0 \end{bmatrix}, \; Y_2 = \begin{bmatrix} A_{m-1} & -I_n & \ldots & 0 \\ A_{m-2} & 0 & \ldots & 0 \\ \vdots & \vdots & & -I_n \\ A_0 & 0 & \ldots & 0 \end{bmatrix}.$$

Both companion forms preserve algebraic and partial multiplicities of all finite eigenvalues of P.

It is shown in [67] that potential linearizations of a matrix polynomial form a linear space. Linearization of a matrix polynomial invariably increases the sensitivity of the eigenvalues. It is therefore important to identify a potentially 'good' linearization which is as well conditioned as possible by analyzing condition numbers of eigenvalues and backward errors of approximate eigenenvalues of linearizations of matrix polynomials. This issue has been investigated in [39, 41].

It is well known that often the eigenvalues of a structured matrix polynomial inherit certain spectral symmetry and that the spectral symmetry often has some physical significance [22, 44, 68, 82, 83]. Therefore, the first step towards solving a structured polynomial eigenvalue problem is to linearize the polynomial in such a way that the linearization reflects the structure of the polynomial and preserves the spectral symmetry. It is shown in [40, 64, 68] that the set of potential structured linearizations of a structured polynomial is an infinite subset of potential linearizations of the polynomial.

The availability of plenty of structured linearizations of a structured matrix polynomial poses a genuine problem of choosing one linearization over other. Since each linearization is expected to be sensitive to perturbations in its own way, for computational purposes, it is highly desirable to identify linearizations of a matrix polynomial which are as well-conditioned as the matrix polynomial itself. This naturally raises some fundamental questions: How does the choice of a linearization affect the accuracy of computed eigenvalues? What distinguishes one linearization from another? Does there exist an optimal linearization? How to identify an optimal or a near optimal linearization? These are certainly fundamental issues which strongly influence numerics of structured polynomial eigenvalue problem.

With a view to answering these questions, we undertake a detailed sensitivity and backward perturbation analysis of structured matrix polynomials and their structured linearizations. We show that structured sensitivity analysis and structured backward perturbation analysis are not only important for accuracy assessment of computed eigenvalues but also play a crucial role in answering the fundamental questions raised above. We mention that the most commonly used stable algorithms such as the QZ algorithm do not preserve structure and hence the spectral-symmetry in the computed eigenvalues [99]. The development of structured preserving algorithms for structured eigenvalue problem is an active and challenging research area [9, 10, 18, 44, 46, 71, 74, 82]. We mention that structured backward perturbation analysis has an important role to play in analyzing stability of structured preserving algorithms.

The sensitivity of a simple eigenvalue of a matrix polynomial is measured by its condition number. An explicit expression for condition number of a simple eigenvalue of a matrix polynomial has been obtained in [1, 93]. With a view to analyzing sensitivity of eigenvalues of

TH-789_04612301

structured matrix polynomials, we define structured condition number of a simple eigenvalue and determine the structured condition number for various structured matrix polynomials. Thus given a simple eigenvalue λ of a structured matrix polynomial P, we determine the structured condition number $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$. Next, we consider a structured linearization L of P and derive the structured condition number $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)$. Further, we identify a potential structured linearization L of P that minimizes $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)/\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$ or $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)/\kappa_{\mathrm{P}}(\lambda)$.

As mentioned before, backward errors play an important role in the accuracy assessment of approximate eigenelements. With a view to analyzing accuracy of approximate eigenelements of a structured eigenvalue problems, we consider structured backward errors of approximate eigenelements. Given an approximate eigenpair (λ, x) of a structured matrix polynomial P, we determine the structured backward error $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Next, we consider a structured linearization L of P and derive the structured backward error $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L})$ of the approximate eigenpair $(\lambda, \Lambda_{m-1} \otimes x)$ of L, where $\Lambda_{m-1} := [\lambda^{m-1}, \ldots, \lambda, 1]^T$. Further, we identify a potential structured linearization L of P that minimizes $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L})/\eta(\lambda, x, \mathrm{P})$.

We show that both the approaches are compatible with each other and lead to the same choice of structured linearization. As structured backward errors together with structured condition numbers provide first order error bounds on approximate eigenelements, these results are of fundamental importance for the numerics of structured polynomial eigenvalue problems.

Pseudospectra of matrices and matrix polynomials provide a powerful framework for analyzing numerics of matrices and matrix polynomials. Pseudospectra of matrices and matrix polynomials have been studied extensively over the years (see, for example, [2, 100] and the references therein). However, for analyzing structured eigenvalue problems, it is necessary to consider structured pseudospectra. We define structured backward error of an approximate eigenvalue of a structured matrix polynomial and analyze structured pseudospectra. We establish a partial equality between structured and unstructured pseudospectra and show that such partial equality plays an important role in solving certain structured distance problems.

Structured mapping problem naturally occurs while addressing some of the issues discussed above. Given two matrices X and B of same size the structured mapping problem requires finding a "structured" square matrix A, if any, such that $AX = B$. Indeed, the task is to find an A that has smallest norm. We produce a complete solution to the structured mapping problem. As a consequence, we determine structured backward errors of approximate invariant subspace of structured matrices. More generally, we consider the structured inverse least-square problem in which the task is to solve the minimization problem

$$\min_A \|AX - B\|_F,$$

where the minimization is taken over structured matrices. we also provide a complete solution of structured inverse least-square problem.

1.2 Preliminaries

In this section we present some basic definitions and results which will be used throughout the thesis. We use standard notations such as $\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$ to denote the vector space of n-tuples $[x_1, \ldots, x_n]^T, x_i \in \mathbb{C}$, and the vector space of m-by-n matrices with real or complex

TH-789_04612301

entries. We denote by A^T and A^H the transpose and conjugate transpose of a matrix $A \in \mathbb{C}^{n\times n}$ respectively. For $A \in \mathbb{C}^{n\times n}$, we denote the spectrum of A by $\sigma(A)$ and is given by $\sigma(A) := \{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\}$.

The singular value decomposition (SVD) of a matrix $A \in \mathbb{C}^{m\times n}$ is given by $A = U\Sigma V^H$, where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary and $\Sigma \in \mathbb{C}^{m\times n}$ is a diagonal matrix with nonnegative diagonal entries (appear in descending order of magnitude). We denote the smallest nonzero singular value of a matrix A by $\sigma_{\min}(A)$.

Now we define the *Kronecker* product and *Hadamard* product of matrices which will be used in the sequel.

Definition 1.2.1. *([84]) Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$. Then the Kronecker product of A and B denoted by $A \otimes B$ is given by the mp-by-nq block matrix*

$$A \otimes B := \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}.$$

Properties of the Kronecker product:([84])

1. Let $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{r\times s}$, $C \in \mathbb{C}^{n\times p}$ and $D \in \mathbb{C}^{s\times t}$. Then
$$(A \otimes B)(C \otimes D) = AC \otimes BD\,(\in \mathbb{C}^{mr\times pt}).$$
2. For all A and B, $(A \otimes B)^T = A^T \otimes B^T$, $\overline{(A \otimes B)} = \overline{A} \otimes \overline{B}$, $(A \otimes B)^H = A^H \otimes B^H$, where $\overline{A}$ denotes the conjugate of A.
3. If A and B are nonsingular, $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
4. For all A and B, $\operatorname{rank}(A \otimes B) = (\operatorname{rank}(A))(\operatorname{rank}(B)) = \operatorname{rank}(B \otimes A)$.
5. Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then
   - $\operatorname{Tr}(A \otimes B) = (\operatorname{Tr}(A))(\operatorname{Tr}(B)) = \operatorname{Tr}(B \otimes A)$
   - $\det(A \otimes B) = (\det(A))^m(\det(B))^n = \det(B \otimes A)$,

   where $\operatorname{Tr}(A)$ is the trace of A and $\det(A)$ denotes the determinant of A.

The Hadamard product of matrices is defined as follows.

Definition 1.2.2. ([84]) *Let $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ and $B = (b_{ij}) \in \mathbb{C}^{m\times n}$. Then the Hadamard product of A and B, denoted by $A \circ B$, is defined as the entry-wise product of A and B, that is, the (i, j)th entry of $A \circ B$ is $a_{ij}b_{ij}$.*

Properties of the Hadamard product:([84]) Suppose $A, B, C \in \mathbb{C}^{m\times n}$ and $z \in \mathbb{C}$. Then

1. $A \circ B = B \circ A$.
2. $A \circ (B + C) = (A \circ B) + (A \circ C)$.
3. $A \circ (zB) = z(A \circ B)$.

TH-789_04612301

We now briefly consider vector and matrix norms to be used in the subsequent development.

Definition 1.2.3. *A function $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ is said to be a norm on $\mathbb{C}^n$ (or a vector norm) if it satisfies the following conditions:*

- $\|x\| = 0 \Leftrightarrow x = 0$.
- $\|\alpha x\| = |\alpha|\|x\|$ *for* $\alpha \in \mathbb{C}$ *and* $x \in \mathbb{C}^n$.
- $\|x + y\| \leq \|x\| + \|y\|$ *for* $x, y \in \mathbb{C}^n$.

Let $\|\cdot\|$ be a norm on $\mathbb{C}^n$. Define $\|\cdot\|_d : \mathbb{C}^n \to \mathbb{R}$ by

$$\|y\|_d := \sup\{|y^H x| : x \in \mathbb{C}^n, \ \|x\| = 1\}.$$

Then it is easy to see that $\|\cdot\|_d$ is a norm and is called the dual norm of the norm $\|\cdot\|$. It follows that for $x, y \in \mathbb{C}^n$, we have $|y^H x| \leq \|x\| \, \|y\|_d$.

Definition 1.2.4. *([87]) A norm $\|\cdot\|$ on $\mathbb{C}^n$ is said to be a monotone if $|x| \leq |y| \Rightarrow \|x\| \leq \|y\|$, where $|x| \leq |y|$ means $|x_j| \leq |y_j|$ for $j = 0 : n$.*

Now we consider matrix norm, that is, norm on $\mathbb{C}^{n \times n}$. Let $\|\cdot\|$ be norm on $\mathbb{C}^n$. Define $\|\cdot\| : \mathbb{C}^{n \times n} \to \mathbb{R}$ by

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{C}^n, \ \|x\| = 1\}.$$

Then $\|\cdot\|$ is a norm on $\mathbb{C}^{n \times n}$ and is referred to as the induced operator norm or the subordinate norm. The subordinate norm induced by the 2-norm $\|\cdot\|_2$ on $\mathbb{C}^n$ is referred to as the spectral norm or the 2-norm on $\mathbb{C}^{n \times n}$. We denote the spectral norm on $\mathbb{C}^{n \times n}$ by $\|\cdot\|_2$. Thus

$$\|A\|_2 := \max_{\|x\|_2 = 1} \|Ax\|_2.$$

The Frobenius norm on $\mathbb{C}^{n \times n}$ is denoted by $\|\cdot\|_F$ and is given by $\|A\|_F := (\mathrm{Tr}(A^H A))^{1/2}$. The spectral and the Frobenius norms have the following useful properties

- $\|Ux\|_2 = \|x\|_2$ if $U^H U = I_n$,
- $\|UAV^H\|_{2,F} = \|A\|_{2,F}$ if $U^H U = I = V^H V$,
- $\|AB\|_{2,F} \leq \|A\|_{2,F}\|B\|_{2,F}$, where the notation $\|\cdot\|_{2,F}$ denotes either $\|\cdot\|_2$ or $\|\cdot\|_F$.

1.2.1 Davis-Kahan-Weinberger dilation theorem

Now we state the Davis-Kahan-Weinberger (DKW) dilation theorem which will be used in the subsequent development.

Let $\mathcal{H} := \mathcal{H}_1 \oplus \mathcal{H}_2$ and $\mathcal{K} := \mathcal{K}_1 \oplus \mathcal{K}_2$ be Hilbert spaces and $\mathbf{T} : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator given by $\mathbf{T} := \begin{bmatrix} A & C \\ B & D \end{bmatrix}$. Then $\mathbf{T}$ is called a *dilation* of A. The norm preserving dilation problem states that given A, B, C and a positive number

$$\mu \geq \max\left(\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|, \left\| \begin{bmatrix} A & C \end{bmatrix} \right\| \right), \tag{1}$$

TH-789_04612301

find all possible D such that $\left\| \begin{bmatrix} A & C \\ B & D \end{bmatrix} \right\| \leq \mu$. Complete solution of this problem is provided by the Davis-Kahan-Weinberger (DKW) theorem, see [24]. We state the result in the context of matrices.

Theorem 1.2.5 (Davis-Kahan-Weinberger, [24])**.** *Let A, B, C be given matrices. Then for any positive number μ satisfying (1), there exists D such that $\left\| \begin{bmatrix} A & C \\ B & D \end{bmatrix} \right\|_2 \leq \mu$. Indeed, those D which have this property are exactly those of the form*

$$D = -KA^H L + \mu(I - KK^H)^{1/2} Z (I - L^H L)^{1/2},$$

where $K^H := (\mu^2 \mathrm{I} - A^H A)^{-1/2} B^H$, $L := (\mu^2 \mathrm{I} - AA^H)^{-1/2} C$ and Z is an arbitrary contraction, that is, $\|Z\|_2 \leq 1$.

We mention that in the case when $(\mu^2 I - A^H A)$ is singular, the inverses in K^H and L are replaced by Moore-Penrose pseudo-inverses (see, [76]).

The DKW Theorem 1.2.5 is a landmark result in the dilation theory of Hilbert space operators and has found applications in many different areas. Specifically, solutions of finite dimensional norm preserving dilation problem (as stated above) have found applications in numerical analysis ([23, 33, 76, 77, 102]). The general solution of DKW Theorem provides solution of norm preserving dilation problem for symmetric, skew-symmetric, Hermitian and skew-Hermitian operators. All matrices in this section are assumed to have entries from $\mathbb{K}$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Corollary 1.2.6. *([24]) Given a Hermitian matrix A and a matrix B, set $\mu := \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_2$. Then there exists $D = D^H$ such that $\left\| \begin{bmatrix} A & B^H \\ B & D \end{bmatrix} \right\| = \mu$. Those D which have this property are of the form*

$$D = -KAK^H + \mu(I - KK^H)^{1/2} Z (I - KK^H)^{1/2},$$

where $K := B(\mu^2 I - A^2)^{-1/2}$ and $Z = Z^H$ is an arbitrary contraction.

Now suppose that A is skew-hermitian. Then the norm preserving skew-Hermitian dilation of A is obtained from the DKW Theorem by considering the fact that $A = A^H$ and $C = -B^H$.

Corollary 1.2.7. *Given a skew-Hermitian matrix A and a matrix B, set $\mu := \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_2$. Then there exists $D = -D^H$ such that $\left\| \begin{bmatrix} A & -B^H \\ B & D \end{bmatrix} \right\|_2 = \mu$. Those D which have this property are of the form*

$$D = -KAK^H + \mu(I - KK^H)^{1/2} Z (I - KK^H)^{1/2}$$

with $K := B(\mu^2 I + A^2)^{-1/2}$ and $Z = -Z^H$ is any arbitrary contraction.

Next, suppose that A is symmetric. Then the norm preserving symmetric dilation of A follows from the DKW Theorem.

TH-789_04612301

Corollary 1.2.8. *Given a symmetric matrix A and a matrix B, set $\mu := \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_2$. Then there exists $D = D^T$ such that $\left\| \begin{bmatrix} A & B^T \\ B & D \end{bmatrix} \right\|_2 = \mu$. Those D which have this property are of the form*

$$D = -K\overline{A}K^T + \mu(I - KK^H)^{1/2} Z (I - \overline{K}K^T)^{1/2},$$

where $K := B(\mu^2 I - \overline{A}A)^{-1/2}$ and $Z = Z^T$ is an arbitrary contraction.

Proof: Using that fact $A^T = A$ and $C = B^T$, by the DKW Theorem, we have $K = B(\mu^2 I - \overline{A}A)^{-1/2}$ and $L = (\mu^2 I - A\overline{A})^{-1/2} B^T = K^T$. Hence $I - KK^H = I - B(\mu^2 I - \overline{A}A)^{-1} B^H$ and $I - L^H L = I - \overline{B}(\mu^2 I - A\overline{A})^{-1} B^T = (I - KK^H)^T = (I - \overline{K}K^T)$. Hence the result follows. ■

Finally, when A is skew-symmetric, the norm preserving skew-symmetric dilations of A are given in the following result.

Corollary 1.2.9. *Given a skew-symmetric matrix A and a matrix B, set $\mu := \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_2$. Then there exists $D = -D^T$ such that $\left\| \begin{bmatrix} A & -B^T \\ B & D \end{bmatrix} \right\|_2 = \mu$. Those D which have this property are of the form*

$$D = -K\overline{A}K^T + \mu(I - KK^H)^{1/2} Z (I - \overline{K}K^T)^{1/2},$$

where $K = B(\mu^2 I + \overline{A}A)^{-1/2}$ and $Z = -Z^T$ is an arbitrary contraction.

Proof: Using the fact that $A = -A^T$ and $C = -B^T$, by the DKW Theorem, we have $K = B(\mu^2 I + \overline{A}A)^{-1/2}$ and $L = -(\mu^2 I + A\overline{A})^{-1/2} B^T = -K^T$. Consequently, we have $I - KK^H = I - B(\mu^2 I + \overline{A}A)^{-1} B^H$ and $I - L^H L = I - \overline{B}(\mu^2 I + A\overline{A})^{-1} B^T = (I - KK^H)^T = (I - \overline{K}K^T)$. Hence the result follows. ■

1.2.2 Matrix polynomials and linearizations

Let $\mathbb{P}_m(\mathbb{C}^{n\times n})$ denote the set of matrix polynomials of the form $\mathrm{P}(z) := \sum_{j=0}^m z^j A_j$, where $A_j \in \mathbb{C}^{n\times n}$. Then $\mathbb{P}_m(\mathbb{C}^{n\times n})$ is a linear space. Now we equip $\mathbb{P}_m(\mathbb{C}^{n\times n})$ with a norm and make it a normed linear space. For $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ given by $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$, we consider the Frobenius and the spectral polynomial norms given by

$$\|\mathrm{P}\|_F := \left(\sum_{j=0}^m \|A_j\|_F^2 \right)^{1/2} \text{ and } \|\mathrm{P}\|_2 := \left(\sum_{j=0}^m \|A_j\|_2^2 \right)^{1/2} \tag{1.2}$$

respectively. Then it follows that, for any $\lambda \in \mathbb{C}, \|\mathrm{P}(\lambda)\| \leq \|\mathrm{P}\|_{F,2} \|\Lambda_m\|_2$, where $\Lambda_m = [1, \lambda, \ldots, \lambda^m]^T$.

A matrix polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is said to be regular if $\det(\mathrm{P}(\lambda)) \neq 0$ for some $\lambda \in \mathbb{C}$. The spectrum of a regular polynomial P, denoted by $\sigma(\mathrm{P})$, is given by

$$\sigma(\mathrm{P}) := \{\lambda \in \mathbb{C} : \det(\mathrm{P}(\lambda)) = 0\}. \tag{1.3}$$

TH-789_04612301

It is possible for P to have an infinite eigenvalue whenever the *determinant* of the leading coefficient of the polynomial is zero. The technical device underlying the notion of the eigenvalue ∞ can be resolved by considering the homogeneous polynomials, see [2, 26]. However, an infinite eigenvalue of a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ can also be resolved by considering the reverse polynomial revP of P given by $\mathrm{revP}(z) := z^m \mathrm{P}(1/z)$ for $z \in \mathbb{C}$. Then ∞ is an eigenvalue of P if and only if 0 is an eigenvalue of revP. Note that $\sigma(\mathrm{P})$ is to be considered as a subset of $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$, the one point compactification of $\mathbb{C}$, whenever P has an infinite eigenvalue. Note also that the map $\Psi : \mathbb{C}_\infty \to \mathbb{C}_\infty, z \mapsto z^{-1}$ is continuous. Hence in the space $\mathbb{C}_\infty$ the following spectral mapping holds:

$$\sigma(\mathrm{revP}) = \{1/\lambda \in \mathbb{C}_\infty : \lambda \in \sigma(\mathrm{P})\} = \Psi(\sigma(\mathrm{P})).$$

In fact rev-operator is a linear isomorphism and $\|\mathrm{P}\|_{F,2} = \|\mathrm{revP}\|_{F,2}$ holds, for details see [2].

The polynomial eigenvalue problem is concerned with finding $(\lambda, x, y) \in \mathbb{C} \times \mathbb{C}^n \times \mathbb{C}^n$ such that

$$\mathrm{P}(\lambda)x = 0 \text{ and } y^H \mathrm{P}(\lambda) = 0,$$

where $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$. The nonzero vectors x and y are called right and left eigenvectors, respectively and λ is called the eigenvalue of P. An eigenvalue λ of a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is called simple if λ is a simple root of the scalar polynomial $\det(\mathrm{P}(\lambda))$. It is shown in [1] that, $y^H(\partial \mathrm{P}(\lambda)/\partial\lambda)x \neq 0$, where $\partial \mathrm{P}(\lambda)/\partial\lambda$ is the first derivative of $\mathrm{P}(\lambda)$ with respect to λ, if and only if λ is a simple eigenvalue of P and x, y are the corresponding right and left eigenvectors respectively.

Due to the lack of a genuine polynomial eigensolver, the standard way of solving polynomial eigenvalue problem is to convert the polynomial P into an equivalent linear polynomial. Given a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, it is possible to convert it into an equivalent linear polynomial

$$\mathrm{L}(\lambda) = \lambda X + Y, \;\; X, Y \in \mathbb{C}^{mn\times mn}.$$

The linear polynomial L is called a *linearization* of P.

Definition 1.2.10. ***(Linearization**,[64])* *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$*. A linear polynomial* $\mathrm{L}(\lambda) = \lambda X + Y$ *with* $X, Y \in \mathbb{C}^{mn\times mn}$ *is called a linearization of* P *if there exist unimodular polynomial (a polynomial* $E(\lambda)$ *is called a unimodular if* $\det E(\lambda)$ *is a nonzero constant, independent of* λ*),* $E(\lambda)$ *and* $F(\lambda)$ *such that*

$$E(\lambda)\mathrm{L}(\lambda)F(\lambda) = \begin{bmatrix} \mathrm{P}(\lambda) & 0 \\ 0 & I_{(m-1)n} \end{bmatrix}, \text{ for all } \lambda \in \mathbb{C}.$$

Note that an immediate consequence of the above definition is that $\gamma \det(\mathrm{L}(\lambda)) = \det(\mathrm{P}(\lambda))$ for some nonzero constant γ. Thus L and P have the same spectrum. In practice the most commonly used examples of linearizations are companion forms or companion polynomials, see [64]. Using these linearizations as prototypes, Mackey et al., [64] introduced two linear spaces of easily constructible linearizations, denoted by $\mathbb{L}_1(\mathrm{P})$ and $\mathbb{L}_2(\mathrm{P})$, and provided a systematic procedure to obtain those linearizations.

TH-789_04612301

Definition 1.2.11. ([64]) *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$*. Then define*

$$\begin{aligned} \mathbb{L}_1(\mathrm{P}) &:= \{\mathrm{L}(\lambda) : \mathrm{L}(\lambda).(\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda),\, v \in \mathbb{C}^m\} \\ \mathbb{L}_2(\mathrm{P}) &:= \{\mathrm{L}(\lambda) : (\Lambda_{m-1}^T \otimes I_n).\mathrm{L}(\lambda) = \omega^T \otimes \mathrm{P}(\lambda),\, \omega \in \mathbb{C}^m\} \end{aligned}$$

where $\Lambda_{m-1} := [\lambda^{m-1}, \lambda^{m-2}, \ldots, \lambda, 1]^T$, $\otimes$ *is the Kronecker product,* v *is called the right ansatz vector for* $\mathrm{L}(\lambda) \in \mathbb{L}_1(\mathrm{P})$ *and* ω *is called the left ansatz vector for* $\mathrm{L}(\lambda) \in \mathbb{L}_2(\mathrm{P})$*.*

It is also shown in [64] that $\mathbb{L}_1(\mathrm{P})$ and $\mathbb{L}_2(\mathrm{P})$ are linear spaces over $\mathbb{C}$ with $\dim \mathbb{L}_1(\mathrm{P}) = m(m-1)n^2 + m = \dim \mathbb{L}_2(\mathrm{P})$.

Notice that when if P is regular, then any linearization L for P must also be regular. In particular, a more surprising result proved in [64] is as follows.

Theorem 1.2.12. *Let* P *be a regular matrix polynomial and let* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$*. Then* L *is a linearization for* P *if and only if* L *is a regular pencil.*

For polynomial eigenvalue problem, the key task for linearizing the polynomial is that, the eigenelements should easily be recovered from the computed eigenelements of the linearization. It is shown in [64] that the linearizations belong to the spaces $\mathbb{L}_1(\mathrm{P})$ and $\mathbb{L}_2(\mathrm{P})$ provide a handy machinery to recover the eigenelements of the given polynomial. The following Theorem illustrates the relation between the eigenelements of P and the eigenelements of its linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ or $\mathrm{L} \in \mathbb{L}_2(\mathrm{P})$.

Theorem 1.2.13. (*Eigenvector recovery from pencils in* $\mathbb{L}_1(\mathrm{P})$, $\mathbb{L}_2(\mathrm{P})$, [64]). *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ *be a regular matrix polynomial and let* $\mathrm{L}(\lambda) \in \mathbb{L}_1(\mathrm{P})$ *with non-zero right ansatz vector* v*. Then* $x \in \mathbb{C}^n$ *is an eigenvector for* P *corresponding to the finite eigenvalue* $\lambda \in \mathbb{C}$ *if and only if* $\Lambda_{m-1} \otimes x$ *is an eigenvector for* $\mathrm{L}(\lambda)$ *corresponding to the eigenvalue* λ*. In fact, every right eigenvector of* L *with finite eigenvalue* λ *is of the form* $\Lambda_{m-1} \otimes x$ *for some eigenvector* x *of* P.

Similarly, if $\mathrm{L} \in \mathbb{L}_2(\mathrm{P})$ *with left ansatz vector* ω*, then* $y \in \mathbb{C}^n$ *is a left eigenvector for* P *corresponding to the finite eigenvalue* $\lambda \in \mathbb{C}$ *if and only if* $\overline{\Lambda_{m-1}} \otimes y$ *is a left eigenvector for* $\mathrm{L}(\lambda)$ *corresponding to the eigenvalue* λ*. In fact, every left eigenvector of* L *with finite eigenvalue* λ *is of the form* $\Lambda_{m-1} \otimes x$ *for some eigenvector* x *of* P.

Recall that a vector $x \in \mathbb{C}^n$ is a right (left) eigenvector of a polynomial P with eigenvalue ∞ if and only if x is the right (left) eigenvector of revP with eigenvalue 0.

Theorem 1.2.14. (*Eigenvector recovery at* ∞, [64]). *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$*. Assume that* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$(*resp.*$\mathbb{L}_2(\mathrm{P})$) *with nonzero right (left) ansatz vector* v*. Then* $x \in \mathbb{C}^n$ *is a right (left) eigenvector for* P *with eigenvalue* ∞ *if and only if* $e_1 \otimes x$ *is a right (left) eigenvector for* L *with eigenvalue* ∞*. If, in addition,* P *is regular and* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ (*resp.* $\mathbb{L}_2(\mathrm{P})$) *is a linearization for* P*, the every right (left) eigenvector of* L *with eigenvalue* ∞ *is of the form* $e_1 \otimes x$ *for some right (left) eigenvector* x *of* P *with eigenvalue* ∞.

By Theorem 1.2.13 and Theorem 1.2.14 it follows that, for an eigenvalue λ of a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, it is necessary to construct two linearizations from each of the two spaces $\mathbb{L}_1(\mathrm{P})$ and $\mathbb{L}_2(\mathrm{P})$ to recover the corresponding right and the left eigenvectors respectively. Now the natural question is: can we have a single linearization which produces both the

TH-789_04612301

right and left eigenvectors? Besides, availability of a very large sources of linearizations give rise a problem of choosing a linearization. These problems are overcome by constructing a new space of potential linearizations, called the *double ansatz space* out of the two spaces of linearizations $\mathbb{L}_1(\mathrm{P})$ and $\mathbb{L}_2(\mathrm{P})$.

Definition 1.2.15. (*Double ansatz space,* [*64*]) *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$. *Then define the double ansatz space as*

$$\mathbb{DL}(\mathrm{P}) := \mathbb{L}_1(\mathrm{P}) \cap \mathbb{L}_2(\mathrm{P}). \tag{1.4}$$

Further, a linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is associated with a right ansatz vector $v \in \mathbb{C}^m$ and $\mathrm{L} \in \mathbb{L}_2(\mathrm{P})$ is associated with a left ansatz vector $\omega \in \mathbb{C}^m$. Hence for $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ there should be some relation between v and ω. The following Theorem gives this relation.

Theorem 1.2.16. ([*64*]) *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ *be a matrix polynomial. Then, for vectors* $v = [v_1, v_2, \ldots, v_m]^T$ *and* $\omega = [\omega_1, \omega_2, \ldots, \omega_m]^T$, *there exist an* $mn \times mn$ *matrix pencil* $\mathrm{L}(\lambda) = \lambda X + Y \in \mathbb{DL}(\mathrm{P})$ *that simultaneously satisfies*

$$\mathrm{L}(\lambda).(\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda), \text{ and } (\Lambda_{m-1}^T \otimes I_n).\mathrm{L}(\lambda) = \omega^T \otimes \mathrm{P}(\lambda)$$

if and only if $v = \omega$.

Notice that every pencil in $\mathbb{DL}(\mathrm{P})$ is not a linearization of a given polynomial [64]. Hence for a given pencil $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$, how to decide that it is a linearization of the given polynomial? What is the connection between the linearization condition of a pencil L and the ansatz vector v which defines L? To answer these questions, a scalar polynomial, called v-polynomial is defined in [64].

Definition 1.2.17. (v-*polynomial,* [*64*]) *Let* $v = [v_1, v_2, \ldots, v_m]^T \in \mathbb{C}^m$. *Define the scalar polynomial*

$$\mathsf{p}(x; v) = v_1 x^{m-1} + v_2 x^{m-2} + \ldots + v_{m-1} x + v_m \tag{1.5}$$

referred to as the "v*-polynomial" of the vector* v. *We adopt the convention that* $\mathsf{p}(x; v)$ *has a root at* ∞ *whenever* $v_1 = 0$.

A connection between the linearization condition of any pencil $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ and the ansatz vector v that defines L, is established in [64] using the v-polynomial. Besides, it is also shown that v-polynomial plays an important role in finding the determinant of $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. The following eigenvalue exclusion Theorem describes the linearization condition.

Theorem 1.2.18. (*Eigenvalue exclusion Theorem,* [*64*]) *Suppose that* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ *and* $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ *is a linearization with ansatz vector* v. *Then* L *is a linearization for* P *if and only if no root of the* v*-polynomial is an eigenvalue of* P. (*Note that this statement includes* ∞ *as one of the possible roots of* $\mathsf{p}(x; \mathsf{v})$ *or possible eigenvalue of* P).

Assume that $\mathrm{L}(\lambda) = \lambda X + Y \in \mathbb{L}_1(\mathrm{P})/\mathbb{DL}(\mathrm{P})$ is a linearization of a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ with respect to the normalized right ansatz vector $v \in \mathbb{C}^m$. Then using Theorem 1.2.13 we have the following lemma.

TH-789_04612301

Lemma 1.2.19. *Let* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})/\mathbb{DL}(\mathrm{P})$ *be a linearization of a polynomial* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, *corresponding to the ansatz vector* $v \in \mathbb{C}^m$. *Then we have*

$$\begin{aligned} \|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2 &= \|v\|_2 \, \|\mathrm{P}(\lambda)x\|_2, & (1.6)\\ |(\Lambda_{m-1} \otimes x)^T \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)| &= |\Lambda_{m-1}^T v| \, |x^T \mathrm{P}(\lambda)x|, & (1.7)\\ |(\Lambda_{m-1} \otimes x)^H \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)| &= |\Lambda_{m-1}^H v| \, |x^H \mathrm{P}(\lambda)x|. & (1.8) \end{aligned}$$

Proof: Assume that $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})/\mathbb{DL}(\mathrm{P})$ is a linearization associated to the ansatz vector $v \in \mathbb{C}^m$ of the given polynomial P. Then we have $\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda)$. Postmultiplying both sides by $(1 \otimes x)$ and then applying norm both sides we obtain

$$\begin{aligned} & \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes I_n)(1 \otimes x) = (v \otimes \mathrm{P}(\lambda))(1 \otimes x) \\ \Rightarrow \;& \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x) = v \otimes \mathrm{P}(\lambda)x \\ \Rightarrow \;& \|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2 = \|v \otimes \mathrm{P}(\lambda)x\|_2 = \|v\|_2 \, \|\mathrm{P}(\lambda)x\|_2. \end{aligned}$$

This proves (1.6).

Further, by Theorem 1.2.13 we know that $x \in \mathbb{C}^n$ is a right eigenvector of P corresponding to the eigenvalue λ if and only if $\Lambda_{m-1} \otimes x$ is a right eigenvector of L corresponding to the eigenvalue λ. Therefore we have

$$\begin{aligned} (\Lambda_{m-1} \otimes x)^T \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x) &= (\Lambda_{m-1} \otimes x)^T (v \otimes \mathrm{P}(\lambda)x) \\ &= (\Lambda_{m-1}^T \otimes x^T)(v \otimes \mathrm{P}(\lambda)x) \\ &= \Lambda_{m-1}^T v \otimes x^T \mathrm{P}(\lambda)x = \Lambda_{m-1}^T v \cdot x^T \mathrm{P}(\lambda)x. \end{aligned}$$

Now taking modulus in both sides we obtain (1.7). The proof is similar for (1.8).■

1.2.3 Structured matrices and matrix polynomials

An n-by-n matrix A is said to be structured if its entries depend on less than n^2 parameters. Hermitian and Hamiltonian matrices are well known examples of structured matrices. A $2n$-by-$2n$ real matrix $\mathcal{H}$ given by $\mathcal{H} := \begin{bmatrix} A & B \\ C & -A^T \end{bmatrix}$ is called a Hamiltonian matrix, where $B = B^T, C = C^T$. Hamiltonian matrices occur in many practical application. It is well known that if λ is an eigenvalue of $\mathcal{H}$ then $-\lambda, \overline{\lambda}, -\overline{\lambda}$ are also eigenvalues of $\mathcal{H}$. In other words, the spectrum of $\mathcal{H}$ is symmetric with respect to real and imaginary axis. It is well known that the set of Hamiltonian matrices forms a Lie algebra whereas the set of Hermitian matrices forms a Jordan Algebra. In the thesis we consider more general classes of structured matrices in the setting of Jordan and Lie algebras associated with appropriate scalar products. We proceed as follows.

Let $\mathbb{K}$ denote the field $\mathbb{R}$ or $\mathbb{C}$ and let $M \in \mathbb{K}^{n\times n}$ be unitary. Consider the scalar product $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \mapsto \langle x, y\rangle_M$ given by

$$\langle x, y\rangle_M := \begin{cases} y^T M x, & \text{bilinear} \\ y^H M x, & \text{sesquilinear,} \end{cases} \qquad (1.9)$$

TH-789_04612301

where x^H denotes the conjugate transpose of x. Then for $A \in \mathbb{K}^{n\times n}$ there is a unique matrix A^* called the adjoint of A with respect to the scalar product $\langle \cdot, \cdot \rangle_M$ such that $\langle Ax, y\rangle_M = \langle x, A^*y\rangle_M$ for all x and y in $\mathbb{K}^n$. An explicit formula for the adjoint is given by

$$A^* = \begin{cases} M^{-1}A^T M, & \text{bilinear form,} \\ M^{-1}A^H M, & \text{sesquilinear form.} \end{cases} \tag{1.10}$$

This gives us two classes of structured matrices, namely, the **Lie algebra**

$$\mathbb{L} := \{A \in \mathbb{K}^{n\times n} : A^* = -A\} \tag{1.11}$$

and the **Jordan algebra**

$$\mathbb{J} := \{A \in \mathbb{K}^{n\times n} : A^* = A\}. \tag{1.12}$$

Further, for bilinear form, we assume that $M^T = \pm M$ and, for sesquilinear form, we assume $M^H = \pm M$. We say that a matrix A is *structured* if A belongs to $\mathbb{J}$ or $\mathbb{L}$. The Jordan and Lie algebras so considered encompass a wide classes of structured matrices. A few examples of structured matrices associated with the scalar products $\langle ., .\rangle_M$ are given in Table 1.1, see [71].

$\mathbb{K}$	M	$\mathbb{J}$	$\mathbb{L}$
$\mathbb{R}^n$	I_n	Symmetric	Skewsymmetric
$\mathbb{C}^n$	I_n	Cplx Symmetric	Cplx Skewsymmetric
$\mathbb{R}^{2n}$	J	Skew-Hamiltonian	Hamiltonian
$\mathbb{C}^{2n}$	J	Cplx J-Skew-symmetric	Cplx J-Symmetric
$\mathbb{C}^n$	I_n	Hermitian	Skew-Hermitian
$\mathbb{C}^n$	$S_{p,q}$	Pseudo-Hermitian	Pseudo-skew-Hermitian
$\mathbb{C}^{2n}$	J	Cplx skew-Hamiltonian	Cplx Hamiltonian

Table 1.1: The Jordan and Lie algebras corresponding to various scalar products.

Here $S_{p,q} := \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$ is the signature matrix and $J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$.

Let $* \in \{T, H\}$. Then observe that the following holds.

$$\begin{aligned} M^* = M : A \in \mathbb{J} \quad &\Leftrightarrow \quad (MA)^* = MA \\ A \in \mathbb{L} \quad &\Leftrightarrow \quad (MA)^* = -MA \end{aligned}$$

$$\begin{aligned} M^* = -M : A \in \mathbb{J} \quad &\Leftrightarrow \quad (MA)^* = -MA \\ A \in \mathbb{L} \quad &\Leftrightarrow \quad (MA)^* = MA. \end{aligned}$$

Obviously, structured matrices can be used to define structured matrix polynomials. For example, a matrix polynomial P is obviously *structured* if the coefficients of P are either in $\mathbb{J}$ or $\mathbb{L}$. However, there are interesting structured matrix polynomials whose coefficient matrices are not the elements of Jordan and/or Lie algebras.

Much of the recent research on structured polynomial eigenvalue problems was motivated

TH-789_04612301

by the second-order polynomial of the form

$$A_0 + zA_1 + z^2 A_0^T, \tag{1.13}$$

where A_1 is complex symmetric: $A_1^T = A_1$. This type of polynomials arise in the study of rail traffic noise caused by high speed trains, see [64, 68, 82]. Notice that the matrix polynomial (1.13) has the property that reversing the order of the coefficients A_0, A_1, followed by taking their transpose, gives back the original polynomial. By analogy with linguistic palindromes, this type of polynomials are called *T-palindromic* polynomials, for further details, see [64, 68, 82].

Consider the *reversal* of a polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$. Then note that $\mathrm{revP}(\lambda)^T = \mathrm{P}(\lambda)^T, \forall\lambda \in \mathbb{C}$ for palindromic polynomials. Similarly if $\mathrm{revP}(\lambda)^T = -\mathrm{P}(\lambda)^T, \forall\lambda \in \mathbb{C}$ it introduces a new structured polynomials which behaves like "*anti*" to the palindromic polynomials and hence this type of polynomials are called *T-anti-palindromic* polynomials.

Further, in gyroscopic systems the following type of polynomial eigenvalue problems occur in the modelling of physical problems. For example

$$(A_0 + \lambda A_1 + \lambda^2 A_2)x = 0, \tag{1.14}$$

where A_0, A_2 are symmetric matrices and A_1 is skew-symmetric matrix: $A_0^T = A_0, A_2^T = A_2, A_1^T = -A_1$. Observe that, replacing λ by $-\lambda$ we obtain the same polynomial eigenvalue problem. Thus $\lambda \mapsto -\lambda$ acts like an even function and this type of polynomials are called *T-even* polynomials. Similarly if $\lambda \mapsto -\lambda$ acts like an odd function then we turn into a new type of structured polynomials called *T-odd* polynomials, for example

$$A_0 + zA_1 + z^2 A_2 \tag{1.15}$$

where A_0, A_2 are skew-symmetric matrices and A_1 is a symmetric matrix: $A_0^T = -A_0, A_2^T = -A_2, A_1^T = A_1$. These two types of polynomials are also related by their reversals. In fact the reversal of a *T-even* polynomial is a *T-odd* polynomial and vice versa.

Its evident that replacing transpose "T" by conjugate transpose "H" of the coefficient matrices of the above types of polynomials we obtain new structured polynomials called *H-palindromic*, *H-anti-palindromic*, *H-even* and *H-odd* polynomials, see [64].

Due to the structure of the coefficient matrices of the polynomial it enforces a symmetry in the spectrum of the structured polynomial discussed above. In the thesis we consider the structured polynomials described in Table 1.2. We denote the space of structured matrix polynomials by $\mathbb{S}$.

In order to consider more general classes of structured polynomials we could replace transpose or conjugate transpose of matrices by adjoint of matrices defined in (1.10), associated with the scalar product $\langle \cdot, \cdot \rangle_M$ defined in (1.9). Thus we could obtain wide classes of structured polynomials identical to that of the polynomials given in Table 1.2.

Structured linearizations of structured matrix polynomials. As mentioned before, the standard approach to solving a polynomial eigenvalue problem proceed by linearizing the matrix polynomial into a matrix pencil of larger size. A question of practical importance

TH-789_04612301

	$P(\lambda) = \sum_{j=0}^{m} \lambda^j A_j \in \mathbb{S}$	
Structure	**Condition**	$\mathbf{m = 2}$
T-symmetric	$P^T(\lambda) = P(\lambda)$	$P(\lambda) = \lambda^2 A_0 + \lambda A_1 + A_2,$ $A_0^T = A_0,\ A_1^T = A_1,\ A_2^T = A_2$
T-skew-symmetric	$P^T(\lambda) = -P(\lambda)$	$P(\lambda) = \lambda^2 A_0 + \lambda A_1 + A_2,$ $A_0^T = -A_0,\ A_1^T = -A_1,\ A_2^T = -A_2$
H-Hermitian	$P^H(\lambda) = P(\bar{\lambda})$	$P(\lambda) = \lambda^2 A_0 + \lambda A_1 + A_2,$ $A_0^H = A_0,\ A_1^H = A_1,\ A_2^H = A_2$
H-skew-Hermitian	$P^H(\lambda) = -P(\bar{\lambda})$	$P(\lambda) = \lambda^2 A_0 + \lambda A_1 + A_2,$ $A_0^H = -A_0,\ A_1^H = -A_1,\ A_2^H = -A_2$
$*$-even	$P^*(\lambda) = P(-\lambda)$	$P(\lambda) = \lambda^2 A + \lambda B + C,$ $A^* = A, B^* = -B, C^* = C$
$*$-odd	$P^*(\lambda) = -P(-\lambda)$	$P(\lambda) = \lambda^2 A + \lambda B + C,$ $A^* = -A, B^* = B, C^* = -C$
$*$-palindromic	$P^*(\lambda) = \lambda^m P(1/\lambda)$	$P(\lambda) = \lambda^2 A + \lambda B + A^*, B^* = B$
$*$-anti-palindromic	$P^*(\lambda) = \lambda^m P(-1/\lambda)$	$P(\lambda) = \lambda^2 A + \lambda B - A^*, B^* = -B$

Table 1.2: Overview of structured matrix polynomials. Note that $* \in \{T, H\}$ may denote either the complex transpose ($* = T$) or the Hermitian transpose ($* = H$).

is whether structured polynomial can be linearized into a matrix pencil which reflects the structure of the polynomial. This issue is investigated in an exhaustive way in [64]. For all the structured polynomials defined above it is showed in [64] that a special class of linearizations that reflect the structure of the polynomial always exist. Those linearizations are called structured linearizations of the structured polynomial. In fact a systematic procedure to construct such linearizations is also discussed in [64]. Define matrices $R, \Sigma \in \mathbb{R}^{m \times m}$ by

$$R = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}, \quad \text{and} \quad \Sigma = \operatorname{diag}\{(-1)^{m-1}, (-1)^{m-2}, \ldots, (-1)^0\}. \tag{1.16}$$

Table 1.3 summarizes the conditions the ansatz vector should satisfy for this purpose(see details in [64]).

structure of P	structure of L	ansatz vector
T-symmetric	T-symmetric	$v \in \mathbb{C}^m$
T-skew-symmetric	T-skew-symmetric	$v \in \mathbb{C}^m$
H-Hermitian	H-Hermitian	$v \in \mathbb{R}^m$
	H-skew-Hermitian	$v \in i\mathbb{R}^m$
H-skew-Hermitian	H-Hermitian	$v \in i\mathbb{R}^m$
	H-skew-Hermitian	$v \in \mathbb{R}^m$
$*$-even	$*$-even	$\Sigma v = (v^*)^T$
	$*$-odd	$\Sigma v = -(v^*)^T$
$*$-odd	$*$-even	$\Sigma v = -(v^*)^T$
	$*$-odd	$\Sigma v = (v^*)^T$
$*$-palindromic	$*$-palindromic	$Rv = (v^*)^T$
	$*$-antipalindromic	$Rv = -(v^*)^T$
$*$-antipalindromic	$*$-palindromic	$Rv = -(v^*)^T$
	$*$-antipalindromic	$Rv = (v^*)^T$

Table 1.3: Structured linearizations of structured polynomials. Here $* \in \{T, H\}$.

TH-789_04612301

Chapter 2

Structured mapping problem

Given two matrices X and B of same size, the structured mapping problem requires to find a "structured" square matrix A such that $AX = B$ and that A has the smallest norm. In this chapter, we investigate structured mapping problem for variety of linearly structured matrices and provide a complete solution.

2.1 Introduction

Let $\mathbb{S}$ be a subset of $\mathbb{K}^{n\times n}$, where $\mathbb{K} := \mathbb{R}$ or $\mathbb{C}$. If $A \in \mathbb{S}$ then we say that the matrix A has structure $\mathbb{S}$. For $X \in \mathbb{K}^{n\times k}$ and $B \in \mathbb{K}^{n\times k}$, consider $\mathbb{S}(X, B) := \{A \in \mathbb{S} : AX = B\}$. The structured mapping problem can be stated as follows.

Problem. Provide a necessary and sufficient condition for $\mathbb{S}(X, B)$ to be nonempty. When $\mathbb{S}(X, B) \neq \emptyset$, characterize elements of $\mathbb{S}(X, B)$ and determine $\inf\{\|Z\| : Z \in \mathbb{S}(X, B)\}$. Further, determine all $A \in \mathbb{S}(X, B)$ such that $\|A\| = \inf\{\|Z\| : Z \in \mathbb{S}(X, B)\}$.

In this chapter, we consider $\mathbb{S} = \mathbb{J}$ or $\mathbb{S} = \mathbb{L}$, where $\mathbb{J}$ and $\mathbb{L}$ are the Jordan and Lie algebra, respectively, associate with appropriate scalar product defined on $\mathbb{K}^n$.

For the special case when $x \in \mathbb{K}^n$, $b \in \mathbb{K}^m$ and $\mathbb{S} = \mathbb{K}^{n\times n}$, Trenkler [101] revisited the mapping problem and has provided solution of the form

$$A = bx^\dagger + Z(I_n - xx^\dagger),$$

where I_n is the $n \times n$ identity matrix, $Z \in \mathbb{K}^{m\times n}$ is arbitrary and $x^\dagger$ is Moore-Penrose inverse of x. Recently, a complete solution of the structured mapping problem for the case when $x \in \mathbb{K}^n, b \in \mathbb{K}^n$ and $\mathbb{S} = \mathbb{J}$ or $\mathbb{S} = \mathbb{L}$ has been provided in [70], where $\mathbb{S}$ is an appropriate Jordan or Lie algebra. Also structured mapping for Hermitian matrices has been analyzed in [91].

In this chapter, we consider $\mathbb{S} = \mathbb{J}$ as well as $\mathbb{S} = \mathbb{L}$ and provide a necessary and sufficient condition for $\mathbb{S}(X, B)$ to be nonempty. Further, we determine each matrix in $\mathbb{S}(X, B)$ and solve the minimization problem

$$\min\{\|A\| : AX = B \text{ and } A \in \mathbb{S}(X, B)\}.$$

TH-789_04612301

We show that the results derived in [70] follow as special cases. As an application of the structured mapping problem, we obtain structured backward errors of approximate eigenelements and approximate invariant subspaces of structured matrices. More specifically, for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $A \in \mathbb{S}$, we define the structured backward error $\eta^{\mathbb{S}}(\lambda, x)$ by

$$\eta^{\mathbb{S}}(\lambda, x) := \inf_{\triangle A \in \mathbb{S}} \{\|\triangle A\| : (A + \triangle A)x = \lambda x\}$$

and determine $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)x = \lambda x$ and $\|\triangle A\| = \eta^{\mathbb{S}}(\lambda, x)$. Further, we define the structured backward error of an approximate invariant subspace $\mathcal{X}$ of A by

$$\eta^{\mathbb{S}}(\mathcal{X}, A) := \inf\{\|\triangle A\| : (A + \triangle A)\mathcal{X} \subset \mathcal{X}\}$$

and determine $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)\mathcal{X} \subset \mathcal{X}$ and $\|\triangle A\| = \eta^{\mathbb{S}}(\mathcal{X}, A)$.

As we shall see, $\mathbb{S}(X, B)$ may be empty for some X and B. In such a case consider the following structured inverse least squared problem (SILSP).

Problem. Solve $\min\{\|AX - B\|_F : A \in \mathbb{S}\} =: \rho_F$ and determine all those $A \in \mathbb{S}$ such that $\|AX - B\|_F = \rho_F$.

Obviously, SILSP is a generalization of the structured mapping problem. SILSP occurs in many applications such as in particle physics and geology, inverse problems of vibration theory, inverse Sturm-Liouville problem, control theory and multidimensional approximation. SILSP has been studied in [90, 89, 106, 107]. We provide a complete solution of the structured inverse least squared problem for the case when $\mathbb{S} = \mathbb{J}$ or $\mathbb{S} = \mathbb{L}$.

Finally, we analyze structured pseudospectra of structured matrices in $\mathbb{S}$. Structured pseudospectra for various structured matrices have been investigated, for example, in [35, 81, 100]. The structured and unstructured pseudospectra of $A \in \mathbb{S}$ are given by

$$\sigma_\epsilon^{\mathbb{S}}(A) := \bigcup_{\|\triangle A\| \le \epsilon} \{\sigma(A + \triangle A) : \triangle A \in \mathbb{S}\} \text{ and } \sigma_\epsilon(A) := \bigcup_{\|\triangle A\| \le \epsilon} \{\sigma(A + \triangle A) : \triangle A \in \mathbb{C}^{n \times n}\},$$

respectively, where $\sigma(A)$ is the spectrum of A. We obtain partial equality between unstructured and structured pseudospectra when $\mathbb{S} = \mathbb{L}$ or $\mathbb{S} = \mathbb{J}$.

2.2 Solution of structured mapping problem

To begin with, we consider structured mapping problem for the four classes of structured matrices, namely, Hermitian, skew-Hermitian, symmetric and skew-symmetric. As we have already seen, these structures are prototypes of more general structures provided by Jordan and Lie algebras associated with appropriate scalar products on $\mathbb{K}^{n \times n}$.

For $A \in \mathbb{K}^{n \times n}$, we denote by A^T the transpose of A and by A^H the conjugate transpose of A. Also whenever necessary, for $* = T$ or $* = H$, we write A^* to denote the transpose or the conjugate transpose of A. Define the map $\mathcal{F} : \mathbb{K}^{n \times k} \times \mathbb{K}^{n \times k} \to \mathbb{K}^{n \times n}$ by

$$\mathcal{F}(X, B) = \begin{cases} BX^\dagger + (BX^\dagger)^T - (X^\dagger)^T X^T B X^\dagger, & \text{if } (X^T B)^T = X^T B \\ BX^\dagger - (BX^\dagger)^T + (X^\dagger)^T X^T B X^\dagger, & \text{if } (X^T B)^T = -X^T B \\ BX^\dagger, & \text{else} \end{cases}$$

TH-789_04612301

where $X^\dagger$ is the Moore-Penrose pseudo-inverse of X. It is easily seen that $\mathcal{F}(X,B)$ is symmetric (resp., skew-symmetric) whenever X^TB is symmetric (resp., skew-symmetric). Also $\mathcal{F}(X,B)X = B$ whenever $BX^\dagger X = B$.

Next define the map $\mathcal{G} : \mathbb{K}^{n\times k} \times \mathbb{K}^{n\times k} \to \mathbb{K}^{n\times n}$ by

$$\mathcal{G}(X,B) = \begin{cases} BX^\dagger + (BX^\dagger)^H - (X^\dagger)^H X^H BX^\dagger, & \text{if } (X^HB)^H = X^HB \\ BX^\dagger - (BX^\dagger)^H + (X^\dagger)^H X^H BX^\dagger, & \text{if } (X^HB)^H = -X^HB \\ BX^\dagger, & \text{else.} \end{cases}$$

Again, it is easily seen that $\mathcal{G}(X,B)$ is Hermitian (resp., skew-Hermitian) whenever X^HB is Hermitian (resp., skew-Hermitian). Also $\mathcal{G}(X,B)X = B$ whenever $BX^\dagger X = B$.

This shows that $\mathcal{F}(X,B)$ and $\mathcal{G}(X,B)$ are potential candidates for solutions of structured mapping problem when $(X^*B)^* = X^*B$ or $(X^*B)^* = -X^*B$ according as the structure under consideration is Hermitian/symmetric or shew-Hermitian/skew-symmetric, where $* \in \{T, H\}$. The following result provides a necessary and sufficient condition for existence of a structured mapping by showing that there is a matrix $A \in \mathbb{K}^{n\times n}$ such that $AX = B$ and $A^* = A$ (resp., $A^* = -A$) if and only if $(X^*B)^* = X^*B$ (resp., $(X^*B)^* = -X^*B$)) and $BX^\dagger X = B$.

To make the presentation simple we use the following notions throughout the rest of the paper.

$$\begin{aligned} \mathsf{sym} &= \{A \in \mathbb{K}^{n\times n} : A^T = A\}, \ \mathsf{skew\text{-}sym} = \{A \in \mathbb{K}^{n\times n} : A^T = -A\}, \ \mathbb{K} := \mathbb{R}/\mathbb{C} \\ \mathsf{Herm} &= \{A \in \mathbb{C}^{n\times n} : A^H = A\}, \ \mathsf{skew\text{-}Herm} = \{A \in \mathbb{C}^{n\times n} : A^H = -A\}. \end{aligned}$$

Theorem 2.2.1. *Let $(X,B) \in \mathbb{K}^{n\times k} \times \mathbb{K}^{n\times k}$. Then there exists a structured matrix $A \in \mathbb{K}^{n\times n}$ if and only if the condition in the Table 2.1 holds.*

A	Condition
sym	$(X^TB)^T = X^TB$ and $BX^\dagger X = B$
skew-sym	$(X^TB)^T = -X^TB$ and $BX^\dagger X = B$
Herm	$(X^HB)^H = X^HB$ and $BX^\dagger X = B$
skew-Herm	$(X^HB)^H = -X^HB$ and $BX^\dagger X = B$

Table 2.1: Necessary and sufficient conditions for structured mapping problem.

In particular, if X has full rank then the condition $BX^\dagger X = B$ is redundant.

Proof: Note that if there exists a structured matrix $A \in \mathbb{K}^{n\times n}$ as specified in the first column of Table 2.1 such that $AX = B$ then obviously the conditions in the second column are satisfied. Indeed, $AX = B$ gives $(X^*B)^* = X^*B$ or $(X^*B)^* = -X^*B$ according as A is symmetric/Hermitian or skew-symmetric/skew-Hermitian, where $* = T$ or $* = H$. Again $AX = B \Rightarrow BX^\dagger X = AXX^\dagger X = AX = B$.

Conversely, if the conditions are satisfied then by constructiuon $A = \mathcal{F}(X,B)$ or $A = \mathcal{G}(X,B)$) gives the desired structured matrix.

Note that if X has full rank then $X^\dagger X = I$ and hence $BX^\dagger X = B$. ■

Let $\mathbb{S}$ denote the set of symmetric/skew-symmetric/Hermitian/skew-hermitian matrices in $\mathbb{K}^{n\times n}$. Often we write this as $\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. Observe that if $A \in$

TH-789_04612301

$\mathbb{K}^{n\times n}$ is given by

$$A := \begin{bmatrix} A_{11} & A_{12}^* \\ A_{12} & A_{22} \end{bmatrix} \text{ then } \|A\|_F = \left(2 \left\| \begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix} \right\|_F^2 - \|A_{11}\|_F^2 + \|A_{22}\|_F^2 \right)^{1/2},$$

where $* = T$ or $* = H$. We repeatedly use this fact in the sequel.

Recall that $\mathbb{S}(X,B) := \{A \in \mathbb{S} : AX = B\}$. Define

$$\eta^{\mathbb{S}}(X,B) := \inf\{\|A\| : A \in \mathbb{S} \text{ and } AX = B\}. \tag{2.1}$$

We denote $\eta^{\mathbb{S}}(X,B)$ by $\eta_2^{\mathbb{S}}(X,B)$ (resp., $\eta_F^{\mathbb{S}}(X,B)$) when $\|\cdot\|$ is the spectral norm (resp., Frobenius norm). The following result provides a complete solution of structured mapping problem for symmetric or skew-symmetric structure.

Theorem 2.2.2. *Let $X, B \in \mathbb{K}^{n\times k}$. Let $\mathbb{S} \subset \mathbb{K}^{n\times n}$ be such that $\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\}$. Suppose that $\operatorname{rank}(X) = r$ and that $\mathbb{S}(X,B) \neq \emptyset$. Then $A \in \mathbb{S}(X,B)$ if and only if A is of the form*

$$A = \mathcal{F}(X,B) + (I - XX^\dagger)^T Z (I - XX^\dagger)$$

for some $Z \in \mathbb{S}$, that is, $\mathbb{S}(X,B) = \mathcal{F}(X,B) + (I - XX^\dagger)^T \mathbb{S}(I - XX^\dagger)$.

Consider the SVD $X := U\Sigma V^H$. Let $U = [U_1\ U_2]$, where $U_1 \in \mathbb{K}^{n\times r}$.

1. **Frobenius norm:** *Then $A := \mathcal{F}(X,B)$ is a unique matrix in $\mathbb{S}$ such that $AX = B$ and*

$$\eta_F^{\mathbb{S}}(X,B) = \|A\|_F = \sqrt{2\|BX^\dagger\|_F^2 - \operatorname{Tr}(BX^\dagger(BX^\dagger)^H(XX^\dagger)^T)}.$$

2. **Spectral norm:** *We have $\eta_2^{\mathbb{S}}(X,B) = \|BX^\dagger\|_2$. Consider the matrix*

$$A = \mathcal{F}(X,B) - (I - XX^\dagger)^T K U_1^H \overline{BX^\dagger U_1} K^T (I - XX^\dagger) + N,$$

where $N = \mu \overline{U}_2 (I - U_2^T K K^H \overline{U}_2)^{1/2} Z (I - U_2^H \overline{K} K^T U_2)^{1/2} U_2^H$, $\mu = \|BX^\dagger\|_2$,

$$K = \begin{cases} BX^\dagger U_1 (\mu^2 I - U_1^H \overline{BX^\dagger} BX^\dagger U_1)^{-1/2}, & \text{if } \mathbb{S} = \mathsf{sym} \\ BX^\dagger U_1 (\mu^2 I + U_1^H \overline{BX^\dagger} BX^\dagger U_1)^{-1/2}, & \text{if } \mathbb{S} = \mathsf{skew\text{-}sym}. \end{cases}$$

and $Z \in \mathbb{S}$ is a contraction. Then $A \in \mathbb{S}$, $AX = B$ and $\|A\|_2 = \eta_2^{\mathbb{S}}(X,B) = \|BX^\dagger\|_2$.

Proof: First, suppose that $\mathbb{S} = \mathsf{sym}$. By assumption there exists $A \in \mathbb{S}$ be such that $AX = B$. Note that $\operatorname{Range}(X) = \operatorname{Range}(U_1)$. Thus representing A relative to the decomposition

$$A : \operatorname{Range}(X) \oplus \operatorname{Range}(X)^\perp \to \operatorname{Range}(\overline{X}) \oplus \operatorname{Range}(\overline{X})^\perp,$$

we have $A = \overline{U}\,\overline{U}^H A U U^H$. Set $\widehat{A} = U^T A U = \begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix}$. Then $\widehat{A} \in \mathbb{S}$ and $\|A\|_{2,F} = \|\widehat{A}\|_{2,F}$. Set $\Sigma_1 := \Sigma(1:r, 1:r)$. Let $V = [V_1, V_2]$ be a conformal partition of V such that $X = U_1 \Sigma_1 V_1$. Now $AX = B \Rightarrow \overline{U}\widehat{A}U^H X = B$. This gives

$$\begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} X = U^T B = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} B \Rightarrow \begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 V_1^H \\ 0 \end{bmatrix} = \begin{bmatrix} U_1^T B \\ U_2^T B \end{bmatrix}$$

TH-789_04612301

$\Rightarrow A_{11}\Sigma_1 V_1^H = U_1^T B$ and $A_{12}\Sigma_1 V_1^H = U_2^T B$.

Therefore, we have $A_{11} = U_1^T BV_1\Sigma_1^{-1} = U_1^T BX^\dagger U_1$, $A_{12} = U_2^T BV_1\Sigma_1^{-1}$. Notice that A_{11} is symmetric if and only if $X^T B = B^T X$ and $BX^\dagger X = B$. Indeed $X^T B = B^T X$ gives $\overline{V}_1\Sigma_1 U_1^T B = B^T U_1\Sigma_1 V_1^H$ and thus $B^T U_1 = \overline{V}_1\Sigma_1 U_1^T BV_1\Sigma_1^{-1}$. Now

$$\begin{aligned}(U_1 BX^\dagger U_1)^T &= U_1^T (X^\dagger)^T B^T U_1 = U_1^T \overline{U}_1 \Sigma_1^{-1} V_1^T B^T U_1 = \Sigma_1^{-1} V_1^T \overline{V}_1 \Sigma_1 U_1^T BV_1\Sigma_1^{-1} \\ &= U_1^T BV_1\Sigma_1^{-1} = U_1^T BX^\dagger U_1\end{aligned}$$

as desired. Thus we have

$$\widehat{A} = \begin{bmatrix} U_1^T BX^\dagger U_1 & (U_2^T BV_1\Sigma_1^{-1})^T \\ U_2^T BV_1\Sigma_1^{-1} & A_{22}\end{bmatrix} \tag{2.2}$$

This shows that $\|\widehat{A}\|_F^2 = 2\|BX^\dagger\|_F^2 - \mathrm{Tr}(BX^\dagger(BX^\dagger)^H(XX^\dagger)^T) + \|A_{22}\|_F^2$. Hence setting $A_{22} = 0$ we obtain a unique matrix

$$A = \overline{U}\begin{bmatrix} U_1^T BX^\dagger U_1 & (U_2^T BV_1\Sigma_1^{-1})^T \\ U_2^T BV_1\Sigma_1^{-1} & 0\end{bmatrix}\begin{bmatrix} U_1^H \\ U_2^H\end{bmatrix} = BX^\dagger + (BX^\dagger)^T - (XX^\dagger)^T BX^\dagger = \mathcal{F}(X,B)$$

such that $A \in \mathbb{S}$, $AX = B$ and $\|A\|_F = \sqrt{2\|BX^\dagger\|_F^2 - \mathrm{Tr}(BX^\dagger(BX^\dagger)^H(XX^\dagger)^T)} = \eta_F^{\mathbb{S}}(X,B)$.

Now from (2.2) we have

$$\begin{aligned} A &= \overline{U}\begin{bmatrix} U_1^T BX^\dagger U_1 & (U_2^T BV_1\Sigma_1^{-1})^T \\ U_2^T BV_1\Sigma_1^{-1} & A_{22}\end{bmatrix}\begin{bmatrix} U_1^H \\ U_2^H\end{bmatrix} \\ &= \begin{bmatrix}\overline{U}_1 & \overline{U}_2\end{bmatrix}\begin{bmatrix} U_1^T BX^\dagger U_1 U_1^H + \Sigma_1^{-1} V_1^T B^T U_2 U_2^T \\ U_2^T BV_1\Sigma_1^{-1} U_1^H + A_{22} U_2^H\end{bmatrix} \\ &= (U_1U_1^H)^T BX^\dagger + (V_1\Sigma_1^{-1}U_1^H)^T B^T(I - U_1U_1^H) + (I - U_1U_1^H)^T BX^\dagger + \overline{U}_2 A_{22} U_2^H \\ &= (XX^\dagger)^T BX^\dagger + (X^\dagger)^T B^T (I - XX^\dagger) + (I - XX^\dagger)^T BX^\dagger + \overline{U}_2 A_{22} U_2^H \\ &= BX^\dagger + (BX^\dagger)^T - (XX^\dagger)^T BX^\dagger + \overline{U}_2 U_2^T \overline{U}_2 A_{22} U_2^H U_2 U_2^H \\ &= BX^\dagger + (BX^\dagger)^T - (X^\dagger)^T X^T BX^\dagger + (I - XX^\dagger)^T Z (I - XX^\dagger) \\ &= \mathcal{F}(X,B) + (I - XX^\dagger)^T Z(I - XX^\dagger),\end{aligned}$$

where $Z \in \mathbb{S}$ is arbitrary.

For the spectral norm, again consider the matrix $\widehat{A}$ given in(2.2) and set $\mu := \left\| \begin{bmatrix} U_1^T BV_1\Sigma_1^{-1} \\ U_2^T BV_1\Sigma_1^{-1}\end{bmatrix}\right\|_2 = \|U^T BV_1\Sigma_1^{-1}\|_2 = \|BX^\dagger\|_2$. Then it follows that $\|\widehat{A}\|_2 \geq \mu$. Now by the DKW Theorem 1.2.8 we have $\|\widehat{A}\|_2 = \mu$ when

$$A_{22} = -K_1\overline{A}_{11}K_1^T + \mu(I - K_1K_1^H)^{1/2} Z (I - \overline{K}_1 K_1^T)^{1/2}$$

where $K_1 = U_2^T BX^\dagger U_1(\mu^2 I - U_1^H \overline{BX^\dagger} BX^\dagger U_1)^{-1/2}$ and $Z \in \mathbb{S}$ is an arbitrary contraction. Thus we have $A = BX^\dagger + (BX^\dagger)^T - (XX^\dagger)^T BX^\dagger + \overline{U}_2 A_{22} U_2^H$ such that $A \in \mathbb{S}$, $AX = B$ and $\|A\|_2 = \|B\mathcal{X}^\dagger\|_2 = \eta_2^{\mathbb{S}}(X,B)$, where A_{22} is defined above. Now simplifying the expression of A we obtain the desired form of A.

The proof is similar for the case when $\mathbb{S} =$ skew-sym. ■

TH-789_04612301

The next result provides a complete solution of structured mapping problem for Hermitian or skew-Hermitian matrices.

Theorem 2.2.3. *Let* $X, B \in \mathbb{K}^{n\times k}$. *Let* $\mathbb{S} \subset \mathbb{K}^{n\times n}$ *be such that* $\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. *Suppose that* $\mathrm{rank}(X) = r$ *and that* $\mathbb{S}(X, B) \neq \emptyset$. *Then* $A \in \mathbb{S}(X, B)$ *if and only if* A *is of the form*

$$A = \mathcal{G}(X, B) + (I - XX^\dagger)Z(I - XX^\dagger)$$

for some $Z \in \mathbb{S}$, *that is,* $\mathbb{S}(X, B) = \mathcal{G}(X, B) + (I - XX^\dagger)\mathbb{S}(I - XX^\dagger)$.

Consider the SVD $X := U\Sigma V^H$. *Let* $U = [U_1\ U_2]$, *where* $U_1 \in \mathbb{K}^{n\times r}$.

1. **Frobenius norm:** *Then* $A := \mathcal{G}(X, B)$ *is a unique matrix in* $\mathbb{S}$ *such that* $AX = B$ *and*
$$\eta_F^{\mathbb{S}}(X, B) = \|A\|_F = \sqrt{2\|BX^\dagger\|_F^2 - \mathrm{Tr}(BX^\dagger(BX^\dagger)^H XX^\dagger)}.$$

2. **Spectral norm:** *We have* $\eta_2^{\mathbb{S}}(X, B) = \|BX^\dagger\|_2$. *Consider the matrix*
$$A = \mathcal{G}(X, B) - (I - XX^\dagger)KU_1^H BX^\dagger U_1 K^H (I - XX^\dagger) + N,$$
where $N := \mu U_2(I - U_2^H KK^H U_2)^{1/2} Z (I - U_2^H KK^H U_2)^{1/2} U_2^H$, $\mu = \|BX^\dagger\|_2$,
$$K = \begin{cases} BX^\dagger U_1(\mu^2 I - U_1^H BX^\dagger BX^\dagger U_1)^{-1/2}, & \text{if } \mathbb{S} = \mathsf{Herm} \\ BX^\dagger U_1(\mu^2 I + U_1^H BX^\dagger BX^\dagger U_1)^{-1/2}, & \text{if } \mathbb{S} = \mathsf{skew\text{-}Herm}. \end{cases}$$
and $Z \in \mathbb{S}$ *is a contraction. Then* $A \in \mathbb{S}$, $AX = B$ *and* $\|A\|_2 = \eta_2^{\mathbb{S}}(X, B) = \|BX^\dagger\|_2$.

Proof: First, suppose that $\mathbb{S} = \mathsf{Herm}$. By assumption there exists $A \in \mathbb{S}$ be such that $AX = B$. Note that $\mathrm{Range}(X) = \mathrm{Range}(U_1)$. Thus representing A relative to the decomposition

$$A : \mathrm{Range}(X) \oplus \mathrm{Range}(X)^\perp \to \mathrm{Range}(X) \oplus \mathrm{Range}(X)^\perp,$$

we have $A = U\ U^H AUU^H$. Set $\widehat{A} = U^H AU = \begin{bmatrix} A_{11} & A_{12}^H \\ A_{12} & A_{22} \end{bmatrix}$. Then $\widehat{A} \in \mathbb{S}$ and $\|A\|_{2,F} = \|\widehat{A}\|_{2,F}$. Set $\Sigma_1 := \Sigma(1 : r, 1 : r)$. Let $V = [V_1, V_2]$ be a conformal partition of V such that $X = U_1\Sigma_1 V_1$. Now $AX = B \Rightarrow U\widehat{A}U^H X = B$. This gives

$$\begin{bmatrix} A_{11} & A_{12}^H \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} X = U^H B = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} B \Rightarrow \begin{bmatrix} A_{11} & A_{12}^H \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 V_1^H \\ 0 \end{bmatrix} = \begin{bmatrix} U_1^H B \\ U_2^H B \end{bmatrix}$$

$\Rightarrow A_{11}\Sigma_1 V_1^H = U_1^H B$ and $A_{12}\Sigma_1 V_1^H = U_2^H B$. Solving these equations, we have

$$\widehat{A} = \begin{bmatrix} U_1^H BX^\dagger U_1 & (U_2^H BV_1\Sigma_1^{-1})^H \\ U_2^H BV_1\Sigma_1^{-1} & A_{22} \end{bmatrix} \tag{2.3}$$

Now we have $\|\widehat{A}\|_F^2 = 2\|BX^\dagger\|_F^2 - \mathrm{Tr}(BX^\dagger(BX^\dagger)^H XX^\dagger) + \|A_{22}\|_F^2$. Hence setting $A_{22} = 0$ we obtain a unique matrix

$$A = U \begin{bmatrix} U_1^H BX^\dagger U_1 & (U_2^H BV_1\Sigma_1^{-1})^H \\ U_2^H BV_1\Sigma_1^{-1} & 0 \end{bmatrix} \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} = BX^\dagger + (BX^\dagger)^H - XX^\dagger BX^\dagger = \mathcal{G}(X, B)$$

TH-789_04612301

such that $A \in \mathbb{S}$, $AX = B$ and $\|A\|_F = \sqrt{2\|BX^\dagger\|_F^2 - \mathrm{Tr}(BX^\dagger(BX^\dagger)^H XX^\dagger)} = \eta_F^{\mathbb{S}}(X, B)$.

Now from (2.3) we have

$$
\begin{aligned}
A &= U \begin{bmatrix} U_1^H BX^\dagger U_1 & (U_2^H BV_1\Sigma_1^{-1})^H \\ U_2^H BV_1\Sigma_1^{-1} & A_{22} \end{bmatrix} \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} \\
&= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} U_1^H BX^\dagger U_1 U_1^H + \Sigma_1^{-1} V_1^H B^H U_2 U_2^H \\ U_2^H BV_1 \Sigma_1^{-1} U_1^H + A_{22} U_2^H \end{bmatrix} \\
&= (U_1U_1^H)BX^\dagger + (V_1\Sigma_1^{-1}U_1^H)^H B^H (I - U_1U_1^H) + (I - U_1U_1^H)BX^\dagger + U_2A_{22}U_2^H \\
&= XX^\dagger BX^\dagger + (X^\dagger)^H B^H (I - XX^\dagger) + (I - XX^\dagger)BX^\dagger + U_2A_{22}U_2^H \\
&= BX^\dagger + (BX^\dagger)^H - (X^\dagger)^H X^H BX^\dagger + (I - XX^\dagger)Z(I - XX^\dagger) \\
&= \mathcal{G}(X, B) + (I - XX^\dagger)Z(I - XX^\dagger),
\end{aligned}
$$

where $Z \in \mathbb{S}$ is arbitrary.

For the spectral norm, again consider the matrix $\widehat{A}$ given in(2.3) and set $\mu := \left\| \begin{bmatrix} U_1^H BV_1\Sigma_1^{-1} \\ U_2^H BV_1\Sigma_1^{-1} \end{bmatrix} \right\|_2 =$ $\|U^H BV_1\Sigma_1^{-1}\|_2 = \|BX^\dagger\|_2$. Then it follows that $\|\widehat{A}\|_2 \geq \mu$. Now by the DKW Theorem 1.2.6 we have $\|\widehat{A}\|_2 = \mu$ when

$$A_{22} = -K_1A_{11}K_1^H + \mu(I - K_1K_1^H)^{1/2} Z (I - K_1K_1^H)^{1/2}$$

where $K_1 = U_2^H BX^\dagger U_1(\mu^2 I - U_1^H BX^\dagger BX^\dagger U_1)^{-1/2}$ and $Z \in \mathbb{S}$ is an arbitrary contraction. Thus we have $A = BX^\dagger + (BX^\dagger)^H - XX^\dagger BX^\dagger + U_2A_{22}U_2^H = \mathcal{G}(X, B) + U_2A_{22}U_2^H$ such that $A \in \mathbb{S}, AX = B$ and $\|A\|_2 = \|BX^\dagger\|_2 = \eta_2^{\mathbb{S}}(X, B)$, where A_{22} is defined above. Now simplifying the expression of A we obtain the desired form of A.

The proof is similar for the case when $\mathbb{S} = \mathsf{skew\text{-}Herm}$. ■

For the special case when X has full rank, the results obtained above have the following simplified forms. Note that when X has full rank in such a case we have $X^\dagger = (X^HX)^{-1}X^H$ and $X^\dagger X = I$. Hence we have the following result whose proof follows from Theorem 2.2.2 and Theorem 2.2.3.

Theorem 2.2.4. *Let* $(X, B) \in \mathbb{K}^{n\times k}$ *and* $\mathbb{S} \subset \mathbb{K}^{n\times n}$ *be such that* $\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. *Suppose that* $\mathrm{rank}(X) = k$ *and that* $\mathbb{S}(X, B) := \{A \in \mathbb{S} : AX = B\} \neq \emptyset$.

1. **Frobenius norm:** *Set* $A := \mathcal{F}(X, B)$ *when* $\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\}$ *and* $A := \mathcal{G}(X, B)$ *when* $\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. *Then A is a unique matrix in $\mathbb{S}$ such that $AX = B$ and*

$$\eta_F^{\mathbb{S}}(X, B) = \|A\|_F = \begin{cases} \sqrt{2\|B(X^HX)^{-1/2}\|_F^2 - \|(X^T\overline{X})^{-1/2}X^TB(X^HX)^{-1/2}\|_F^2}, \\ \qquad \text{if } \mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\} \\ \sqrt{2\|B(X^HX)^{-1/2}\|_F^2 - \|(X^HX)^{-1/2}X^HB(X^HX)^{-1/2}\|_F^2}, \\ \qquad \text{if } \mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}. \end{cases}$$

2. **Spectral norm:** *We have* $\eta_2^{\mathbb{S}}(X, B) = \|B(X^HX)^{-1/2}\|_2$. *For* $* \in \{T, H\}$, *consider* $Q := ((X^HX)^{-1/2})^* X^* B(X^HX)^{-1/2}$ *and* $K := B(X^HX)^{-1/2}(\mu^2 I - Q^HQ)^{-1/2}$, *where* $\mu := \|B(X^HX)^{-1/2}\|_2$. *Let U be an isometry such that* $U^HX = 0$.

TH-789_04612301

When $* = T$, *define*

$$A := \mathcal{F}(X,B) - (I - XX^\dagger)^T K(X^HX)^{-1/2}\overline{X^TB}(K(X^HX)^{-1/2})^T(I-XX^\dagger) + N,$$

where $N = \mu\overline{U}(I - U^TKK^H\overline{U})^{1/2}Z(I - U^H\overline{K}K^TU)^{1/2}U^H$ *and* $Z \in \mathbb{S}$ *is a contraction.*
When $* = H$, *define*

$$A := \mathcal{G}(X,B) - (I - XX^\dagger)K(X^HX)^{-1/2}X^HB(K(X^HX)^{-1/2})^H(I-XX^\dagger) + N,$$

where $N = \mu U(I - U^HKK^HU)^{1/2}Z(I - U^HKK^HU)^{1/2}U^H$ *and* $Z \in \mathbb{S}$ *is a contraction.*
Then $A \in \mathbb{S}$, $AX = B$ *and* $\|A\|_2 = \eta_2^{\mathbb{S}}(X,B) = \|B(X^HX)^{-1/2}\|_2$.

Now we present the solution of structured mapping problem for structured matrices when $\mathbb{S} = \mathbb{J}$ or $\mathbb{S} = \mathbb{L}$. Recall that $\mathbb{J}$ and $\mathbb{L}$ are Jordan and Lie algebras associated with the scalar product $\langle x, y\rangle_M := y^*Mx$, where M is unitary, $M^* = \pm M$ and $* \in \{T, H\}$. Also recall that $A \in J$ if and only if $(MA)^* = MA$ or $(MA)^* = MA$ according as $M^* = M$ or $M^* = -M$. Similarly, $A \in \mathbb{L}$ if and only if $(MA)^* = MA$ or $(MA)^* = -MA$ according as $M^* = -M$ or $M^* = M$. This shows that if $S \in \{\mathbb{J}, \mathbb{L}\}$ then defining $M\mathbb{S} := \{MA : A \in \mathbb{S}\}$ we have $M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. Consequently, for $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$, we immediately obtain the following necessary and sufficient condition for existence of a structured mapping in $\mathbb{S}$.

Theorem 2.2.5. *Let* $(X, B) \in \mathbb{K}^{n\times k} \times \mathbb{K}^{n\times k}$. *Let* $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$. *Then there is a matrix* $A \in \mathbb{S}$ *such that* $AX = B$ *if and only if the conditions in the following table holds.*

M	$\mathbb{J}$	$\mathbb{L}$
$M^T = M$	$(X^TMB)^T = X^TMB$ $(MB)X^\dagger X = MB$	$(X^TMB)^T = -X^TMB$ $(MB)X^\dagger X = MB$
$M^T = -M$	$(X^TMB)^T = -X^TMB$ $(MB)X^\dagger X = MB$	$(X^TMB)^T = X^TMB$ $(MB)X^\dagger X = MB$
$M^H = M$	$(X^HMB)^H = X^HMB$ $(MB)X^\dagger X = MB$	$(X^HMB)^H = -X^HMB$ $(MB)X^\dagger X = MB$
$M^T = -M$	$(X^HMB)^H = -X^HMB$ $(MB)X^\dagger X = MB$	$(X^HMB)^H = X^HMB$ $(MB)X^\dagger X = MB$

For the special case of nonzero vectors $x \in \mathbb{K}^n$ and $b \in \mathbb{K}^n$, by Theorem 2.2.5, we obtain the following necessary and sufficient condition provided in [70]:

Recall that if $S \in \{\mathbb{J}, \mathbb{L}\}$ then $M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$. Therefore, in view of Theorem 2.2.2 and Theorem 2.2.3, we have the following result whose proof is immediate.

Theorem 2.2.6. *Let* $X, B \in \mathbb{K}^{n\times k}$. *Let* $\mathbb{S} \subset \mathbb{K}^{n\times n}$ *be such that* $\mathbb{S} \subset \{\mathbb{J}, \mathbb{L}\}$. *Suppose that* $\mathrm{rank}(X) = r$ *and that* $\mathbb{S}(X, B) \neq \emptyset$. *Let* $A \in \mathbb{S}(X, B)$. *Then* A *is of the form*

$$A = M^{-1}\mathcal{F}(X, MB) + M^{-1}(I - XX^\dagger)^T Z(I - XX^\dagger)$$

TH-789_04612301

M	$\mathbb{J}$	$\mathbb{L}$
$M^T = M$	$x, b \in \mathbb{K}^n$	$\langle x, b\rangle_M = 0$
$M^T = -M$	$x^T M b = 0$	$x, b \in \mathbb{K}^n$
$M^H = M$	$x^H M b \in \mathbb{R}$	$x^H M b \in i\mathbb{R}$
$M^H = -M$	$x^H M b \in i\mathbb{R}$	$x^H M b \in \mathbb{R}$

Table 2.2: Necessary and sufficient condition

for some $Z \in \mathbb{S}$, *that is,* $\mathbb{S}(X, B) = M^{-1}\mathcal{F}(X, MB) + M^{-1}(I - XX^\dagger)^T \mathbb{S}(I - XX^\dagger)$.

Consider the SVD $X := U\Sigma V^H$. *Let* $U = [U_1\ U_2]$, *where* $U_1 \in \mathbb{K}^{n\times r}$.

1. **Frobenius norm:** *When* $M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\}$ *define* $A := M^{-1}\mathcal{F}(X, MB)$ *and when* $M\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$ *define* $A := M^{-1}\mathcal{G}(X, MB)$. *Then* A *is a unique matrix in* $\mathbb{S}$ *such that* $AX = B$ *and*

$$\eta_F^{\mathbb{S}}(X, B) = \|A\|_F = \begin{cases} \sqrt{2\|BX^\dagger\|_F^2 - \mathrm{Tr}(MBX^\dagger(MBX^\dagger)^H(XX^\dagger)^T)}, & \text{if } M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\} \\ \sqrt{2\|BX^\dagger\|_F^2 - \mathrm{Tr}(MBX^\dagger(MBX^\dagger)^H XX^\dagger)}, & \text{if } M\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}. \end{cases}$$

2. **Spectral norm:** *We have* $\eta_2^{\mathbb{S}}(X, B) = \|BX^\dagger\|_2$.

   When $M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\}$ *consider the matrix*

$$A = M^{-1}\mathcal{F}(X, MB) - M^{-1}(I - XX^\dagger)^T K U_1^H \overline{MBX^\dagger U_1} K^T (I - XX^\dagger) + M^{-1}N,$$

   where $N = \mu\overline{U}_2(I - U_2^T K K^H \overline{U}_2)^{1/2} Z (I - U_2^H \overline{K} K^T U_2)^{1/2} U_2^H$, $\mu = \|BX^\dagger\|_2$,

$$K = \begin{cases} MBX^\dagger U_1(\mu^2 I - U_1^H \overline{MBX^\dagger} MBX^\dagger U_1)^{-1/2}, & \text{if } M\mathbb{S} = \mathsf{sym} \\ MBX^\dagger U_1(\mu^2 I + U_1^H \overline{MBX^\dagger} MBX^\dagger U_1)^{-1/2}, & \text{if } M\mathbb{S} = \mathsf{skew\text{-}sym}. \end{cases}$$

   and $Z \in \mathbb{S}$ *is a contraction.*

   When $M\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$ *consider the matrix*

$$A = M^{-1}\mathcal{G}(X, MB) - M^{-1}(I - XX^\dagger) K U_1^H MBX^\dagger U_1 K^H (I - XX^\dagger) + M^{-1}N,$$

   where $N := \mu U_2(I - U_2^H K K^H U_2)^{1/2} Z (I - U_2^H K K^H U_2)^{1/2} U_2^H$, $\mu = \|BX^\dagger\|_2$,

$$K = \begin{cases} MBX^\dagger U_1(\mu^2 I - U_1^H MBX^\dagger MBX^\dagger U_1)^{-1/2}, & \text{if } M\mathbb{S} = \mathsf{Herm} \\ MBX^\dagger U_1(\mu^2 I + U_1^H MBX^\dagger MBX^\dagger U_1)^{-1/2}, & \text{if } M\mathbb{S} = \mathsf{skew\text{-}Herm}. \end{cases}$$

   and $Z \in \mathbb{S}$ *is a contraction. Then* $A \in \mathbb{S}$, $AX = B$ *and* $\|A\|_2 = \eta_2^{\mathbb{S}}(X, B) = \|BX^\dagger\|_2$.

When X has full rank, it is easy to obtain an analogue of Theorem 2.2.4 for structured matrices in $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$.

TH-789_04612301

2.3 Structured backward errors of eigenelements

We now show that structured mapping problem naturally arises when analyzing backward errors of approximate eigenelements and approximate invariant subspaces of structured matrices. Let $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$. For the rest of this section, we assume that $\mathbb{K} = \mathbb{C}$. Let $A \in \mathbb{S}$.

Now, let $X \in \mathbb{C}^{n\times p}$ be a full column rank matrix, that is, $\mathrm{rank}(A) = p$. Let $D \in \mathbb{C}^{p\times p}$. We wish to find a matrix $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)X = XD$ so that X is an invariant subspace of $A + \triangle A$. Indeed, we wish to find $\triangle A$ that has the smallest norm. Observe that such a matrix $\triangle A$ exists if and only if $\triangle A$ satisfies $\triangle AX = R$, where $R := XD - AX$. Thus the problem of finding $\triangle A$ boils down to a solution of structured mapping problem. For the rest of this section, we denote $\eta^{\mathbb{S}}(X, R)$ by $\eta^{\mathbb{S}}(D, X)$. Then

$$\eta^{\mathbb{S}}(D, X) := \inf\{\|\triangle A\| : \triangle A \in \mathbb{S} \text{ and } (A + \triangle A)X = XD\}. \tag{2.4}$$

For the spectral and the Frobenius norms, we denote $\eta^{\mathbb{S}}(X, D)$ by $\eta^{\mathbb{S}}_2(D, X)$ and $\eta^{\mathbb{S}}_F(D, X)$, respectively. Note that Theorem 2.2.5 provides a necessary and sufficient condition for existence of $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)X = XD$. Further, for the spectral and the Frobenius norms, Theorem 2.2.6 provides $\eta^{\mathbb{S}}(D, X)$ and a matrix $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)X = XD$ and $\|\triangle A\| = \eta^{\mathbb{S}}(D, X)$. Since X has full rank, we easily obtain that $\eta^{\mathbb{S}}_2(D, X) = \|R(X^HX)^{-1/2}\|_2$ and

$$\eta^{\mathbb{S}}_F(D, X) = \begin{cases} \sqrt{2\|R(X^HX)^{-1/2}\|_F^2 - \|(X^T\overline{X})^{-1/2}X^TMR(X^HX)^{-1/2}\|_F^2}, \\ \hfill \text{if } M\mathbb{S} \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\} \\ \sqrt{2\|R(X^HX)^{-1/2}\|_F^2 - \|(X^HX)^{-1/2}X^HMR(X^HX)^{-1/2}\|_F^2}, \\ \hfill \text{if } M\mathbb{S} \in \{\mathsf{Herm}, \mathsf{skew\text{-}Herm}\}. \end{cases}$$

We now consider the special case when $X = x \in \mathbb{C}^n$ and $D = \lambda \in \mathbb{C}$ and derive structured backward error $\eta^{\mathbb{S}}(\lambda, x)$. Before we proceed further, we briefly discuss the spectral symmetry of eigenvalues of a structured matrix. Table 2.3 summaries some important spectral symmetries of eigenvalues of structured matrices. These results follow from the fact that when $A \in \mathbb{S}$, we have either $M^{-1}A^TM = A$ or $M^{-1}A^HM = A$.

Recall that (λ, y, x) is an eigentriple of A if $Ax = \lambda x$ and $y^HA = \lambda y^H$. Spectral symmetries of structured matrices are also reflected in their eigentriples. Table 2.3 summarizes eigentriples of the structured matrices.

Scalar product	y^TMx	y^HMx	y^TMx
Field	$\mathbb{K} = \mathbb{C}$	$\mathbb{K} = \mathbb{C}$	$\mathbb{K} = \mathbb{R}$
$\mathbb{J}$	λ	$(\lambda, \overline{\lambda})$	$(\lambda, \overline{\lambda})$
$\mathbb{L}$	$(\lambda, -\lambda)$	$(\lambda, -\overline{\lambda})$	$(\lambda, -\lambda, \overline{\lambda}, -\overline{\lambda})$
$\mathbb{J}$	$(\lambda, \overline{Mx}, x)$	$(\lambda, My, x), (\overline{\lambda}, y, Mx)$	(λ, Mx, x)
$\mathbb{L}$	$(\lambda, \overline{My}, x), (-\lambda, y, \overline{Mx})$	$(\lambda, My, x), (-\overline{\lambda}, y, Mx)$	$(\lambda, My, x), (-\lambda, y, Mx)$

Table 2.3: Spectral symmetry of structured matrices.

Set $r := \lambda x - Ax$. Then there is a matrix $E \in \mathbb{S}$ such that $(A + E)x = \lambda x$ if and only if x

TH-789_04612301

and r satisfy the condition in Table 2.2. Suppose that x and r satisfy the condition. Then it follows that $\eta_2^{\mathbb{S}}(\lambda, x) = \|r\|_2/\|x\|_2$ and $\eta_F^{\mathbb{S}}(\lambda, x) = \sqrt{2\|r\|_2^2 - |\langle r, x\rangle_M|^2} \leq \sqrt{2}\eta_2^{\mathbb{S}}(\lambda, x)$.

Now by Theorem 2.2.6, defining E by

$$E = \begin{cases} (x^T r)M^{-1}\overline{x}x^H + M^{-1}\overline{x}r^T(I - xx^H) + M^{-1}(I - \overline{x}x^T)rx^H, & \text{if } MA \in \mathsf{sym} \\ M^{-1}(I - \overline{x}x^T)rx^H - M^{-1}\overline{x}r^T(I - xx^H), & \text{if } MA \in \mathsf{skew\text{-}sym} \\ (x^H r)M^{-1}xx^H + M^{-1}xr^H(I - xx^H) + M^{-1}(I - xx^H)rx^H, & \text{if } MA \in \mathsf{Herm} \\ (x^H r)M^{-1}xx^H - M^{-1}xr^H(I - xx^H) + M^{-1}(I - xx^H)rx^H, & \text{if } MA \in \mathsf{skew\text{-}Herm}. \end{cases}$$

we have $E \in \mathbb{S}$ such that $(A + E)x = \lambda x$ and $\|E\|_F = \eta_F^{\mathbb{S}}(\lambda, x)$. Further, defining $\triangle A$ by

$$\triangle A = \begin{cases} E - \dfrac{\overline{x^T r}M^{-1}(I - \overline{x}x^T)rr^T(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}, & \text{if } MA \in \mathsf{sym} \\ E, & \text{if } MA \in \mathsf{skew\text{-}sym} \\ E - \dfrac{\overline{x^H r}M^{-1}(I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}, & \text{if } MA \in \mathsf{Herm} \\ E - \dfrac{x^H r M^{-1}(I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}, & \text{if } MA \in \mathsf{skew\text{-}Herm}. \end{cases}$$

we have $\triangle A \in \mathbb{S}$ such that $(A + \triangle A)x = \lambda x$ and $\|\triangle A\|_2 = \eta_2^{\mathbb{S}}(\lambda, x)$.

2.3.1 Structured pseudospectra

Structured pseudospectra provide a convenient framework for analyzing numerics of structured matrices. Let $A \in \mathbb{S}$ be a structured matrix. Then the structured ϵ-pseudospectrum of A, which we denote by $\sigma_\epsilon^{\mathbb{S}}(A)$, is defined by

$$\sigma_\epsilon^{\mathbb{S}}(A) := \bigcup_{\|\triangle A\| \leq \epsilon} \{\sigma(A + \triangle A) : \triangle A \in \mathbb{S}\}$$

where $\sigma(A)$ denotes the spectrum of A. Structured pseudospectra are characterized by structured backward errors of approximate eigenvalues of A. For $A \in \mathbb{S}$ and $z \in \mathbb{C}$, define

$$\eta^{\mathbb{S}}(z, A) := \inf\{\|\triangle A\| : z \in \sigma(A + \triangle A),\ \triangle A \in \mathbb{S}\}.$$

Then it follows that $\eta^{\mathbb{S}}(\lambda, A) = \min_{\|x\|_2=1}\{\eta^{\mathbb{S}}(\lambda, x) : x \in \mathbb{C}^n\}$. We mention that for some z, the infimum may not be attained. In such a case, we set $\eta^{\mathbb{S}}(z, A) = \infty$. Thus $\eta^{\mathbb{S}}(z, A)$ is the structured backward error of z when z is treated as an approximate eigenvalue value of A. For the spectral and the Frobenius norms, we denote $\eta^{\mathbb{S}}(\lambda, A)$ by $\eta_2^{\mathbb{S}}(\lambda, A)$ and $\eta_F^{\mathbb{S}}(\lambda, A)$ respectively.

We set $\eta(\lambda, A) := \sigma_{\min}(A - \lambda I)$. Then it is easily seen that $\eta(\lambda, A)$ is the unstructured backward error of λ as an approximate eigenvalue of A. Unlike in the case of unstructured backward error $\eta(z, A)$, determination of structured backward error $\eta^{\mathbb{S}}(z, A)$ for many interesting structures is an open problem. Consequently, determining $\sigma_\epsilon^{\mathbb{S}}(A)$ for those structures is an open problem. Note that $\eta^{\mathbb{S}}(z, A) \geq \eta(z, A)$. For certain structures and for certain $z \in \mathbb{C}$, it turns that $\eta^{\mathbb{S}}(z, A) = \eta(z, A)$.

Recall that $MA \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$ when $A \in \mathbb{J} \subset \mathbb{C}^{n\times n}$ or $A \in \mathbb{L} \subset \mathbb{C}^{n\times n}$ and $M^* = \pm M, * \in \{T, H\}$.

TH-789_04612301

Theorem 2.3.1. *Let* $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \mathsf{sym}\}$ *when* $M^T = M$, *and* $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}\}\}$ *when* $M^T = -M$. *Let* $A \in \mathbb{S}$. *Then for* $z \in \mathbb{C}$, *there exists* $\triangle A \in \mathbb{S}$ *such that* $z \in \sigma(A + \triangle A)$ *and* $\eta^{\mathbb{S}}(z, A) = \|\triangle A\| = \eta(z, A)$. *Consequently, we have*

$$\sigma_\epsilon^{\mathbb{S}}(A) = \sigma_\epsilon(A).$$

Proof: Consider the case $MA \in \mathsf{sym}$ and $M = M^T$. Let $\lambda \in \mathbb{C}$. It follows that $A - \lambda I \in \mathbb{S}$, that is, $(M(A - \lambda I))^T = M(A - \lambda I)$. Since $M(A - \lambda I)$ is complex symmetric there exists a unitary matrix U such that the symmetric Takagi factorization [45] $M(A-\lambda I) = U\Sigma U^T$ holds, where Σ is a diagonal matrix containing singular values of $M(A - \lambda I)$ ordered in descending order of magnitude. Note that $\eta(\lambda, A) = \sigma_{\min}(A - \lambda I) = \sigma_{\min}(M(A - \lambda I)) = \Sigma(n, n)$. Let $u := U(:, n)$. Then we have $M(A - \lambda I)\overline{u} = \eta(\lambda, A)\overline{u}$. This gives $(A - \eta(\lambda, A)M^{-1}uu^T)\overline{u} = \lambda\overline{u}$. Setting $\triangle A := -\eta(\lambda, A)M^{-1}uu^T$, we have $\triangle A \in \mathbb{S}$ and $\|\triangle A\| = \eta(\lambda, A) = \eta^{\mathbb{S}}(\lambda, A)$. Hence the results follow.

Next, consider the case $MA \in \mathsf{skew\text{-}sym}$ and $M^T = -M$. Then for $\lambda \in \mathbb{C}$, we have $M(A - \lambda I) \in \mathbb{S}$, that is, $(M(A - \lambda I))^T = -M(A - \lambda I)$. Since $M(A - \lambda I)$ is complex skewsymmetric, we have the skewsymmetric Takagi factorization [45]

$$M(A - \lambda I) = U\text{diag}(d_1, \cdots, d_m)U^T,$$

where U is unitary, $d_j := \begin{bmatrix} 0 & s_j \\ -s_j & 0 \end{bmatrix}$, $s_j \in \mathbb{C}$ is nonzero and $|s_j|$ are singular values of $M(A - \lambda I)$. Here the blocks d_j appear in descending order of magnitude of $|s_j|$. Note that $M(A - \lambda I)\overline{U} = U\text{diag}(d_1, \cdots, d_m)$. Let $u := U(:, n-1:n)$. Then $M(A - \lambda I)\overline{u} = ud_m = ud_m u^T\overline{u}$. This gives $(MA - ud_m u^T)\overline{u} = \lambda M\overline{u}$. Hence taking $\triangle A := -M^{-1}ud_m u^T$, we have $\lambda \in \sigma(A + \triangle A)$, $\triangle A \in \mathbb{S}$ and $\|\triangle A\| = |s_m| = \sigma_{\min}(M(A - \lambda I)) = \sigma_{\min}(A - \lambda I) = \eta(\lambda, A)$. Hence $\eta^{\mathbb{S}}(\lambda, A) = \eta(\lambda, A)$ and the desired result follows.

Finally, consider the case $M^T = -M$ and $MA \in \mathsf{sym}$. Let $\lambda \in \mathbb{C}$. Note that $\sigma_{\min}(M(A - \lambda I)) = \sigma_{\min}(A - \lambda I) = \eta(\lambda, A)$. Let u and v be unit left and right singular vector of $M(A-\lambda I)$ corresponding to $\eta(\lambda, A)$. Then $M(A - \lambda I)v = \eta(\lambda, A)u$. This gives $Av - \eta(\lambda, A)M^{-1}u = \lambda v$. Let $E \in \mathbb{C}^{n\times n}$ be such that $E = E^T$, $Ev = u$ and $\|E\| = 1$. Such a matrix always exists. Then setting $\triangle A := -\eta(\lambda, A)M^{-1}E$, we have $(A + \triangle A)v = \lambda v$, $\triangle A \in \mathbb{S}$ and $\|\triangle A\| = \eta(\lambda, A) = \eta^{\mathbb{S}}(\lambda, A)$. Hence the result follows. ■

For Lie and Jordan algebras corresponding to sesquilinear form induced by M, we have partial equality between structured and unstructured pseudospectra.

Theorem 2.3.2. *Let* $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \mathsf{Herm}\}$ *when* $M^H = M$ *and* $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \mathsf{skew\text{-}Herm}\}$ *when* $M^H = -M$. *Let* $A \in \mathbb{S}$. *Then for* $z \in \mathbb{R}$, *there exists* $\triangle A \in \mathbb{S}$ *such that* $z \in \sigma(A + \triangle A)$ *and* $\eta^{\mathbb{S}}(z, A) = \|\triangle A\| = \eta(z, A)$. *Consequently, we have*

$$\sigma_\epsilon^{\mathbb{S}}(A) \cap \mathbb{R} = \sigma_\epsilon(A) \cap \mathbb{R}.$$

Next, consider $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \mathsf{skew\text{-}Herm}\}$ *when* $M^H = M$ *and* $\mathbb{S} := \{A \in \mathbb{C}^{n\times n} : MA \in \mathsf{Herm}\}$ *when* $M^H = -M$. *Let* $A \in \mathbb{S}$. *Then for* $z \in i\mathbb{R}$, *there exists* $\triangle A \in \mathbb{S}$ *such that*

TH-789_04612301

$z \in \sigma(A + \triangle A)$ *and* $\eta^{\mathbb{S}}(z, A) = \|\triangle A\| = \eta(z, A)$. *Consequently, we have*

$$\sigma_{\epsilon}^{\mathbb{S}}(A) \cap i\mathbb{R} = \sigma_{\epsilon}(A) \cap i\mathbb{R}.$$

Proof: Consider the case $MA \in \mathsf{Herm}$ when $M = M^H$. Then for $\lambda \in \mathbb{R}$, $M(A-\lambda I) \in \mathbb{S}$. Since $M(A-\lambda I)$ is hermitian, we have the spectral decomposition $M(A-\lambda I) = U\text{diag}(\mu_1, \cdots, \mu_n)U^H$, where U is unitary and μ_j's appear in descending order of their magnitudes. Note that $|\mu_n| = \sigma_{\min}(M(A - \lambda I)) = \sigma_{\min}(A - \lambda I) = \eta(\lambda, A)$. Now defining $\triangle A := -\mu_n M^{-1}U(:, n)U(:, n)^H$, we have $\lambda \in \sigma(A + \triangle A)$, $\triangle A \in \mathbb{S}$ and $\|\triangle A\| = \eta(\lambda, A) = \eta^{\mathbb{S}}(\lambda, A)$. Hence the result follows. The proof is similar for the case when $MA \in \mathsf{skew\text{-}Herm}$ and $M^H = -M$.

Finally, consider the case when $MA \in \mathsf{skew\text{-}Herm}$ and $M^H = M$. Then for $\lambda \in i\mathbb{R}$, the set of purely imaginary numbers, we have $M(A - \lambda I) \in \mathbb{S}$, that is, $M(A - \lambda I)$ is skew hermitian. Hence the result follows from spectral decomposition of $M(A - \lambda I)$. The proof is similar for the case when $MA \in \mathsf{Herm}$ and $M^H = -M$. ■

Similar results hold for some other important structured matrices such as Toeplitz and Hankel matrices which are not described in the setting of Jordan or Lie algebras.

2.4 Structured backward error for approximate invariant subspaces of structured matrices

In this section we find an explicit expression of structured backward error of approximate invariant subspaces of structured matrices. First we briefly review the relation between the norm of a matrix T with norm of the *pinching* of T, (see [12, 13].)

Let T be a matrix given by $\mathrm{T} := \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. The diagonal matrix $\mathcal{P}(\mathrm{T}) := \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$ is called the *pinching* of T. Then we have

$$\|\mathcal{P}(\mathrm{T})\| \leq \|\mathrm{T}\| \tag{2.5}$$

for any arbitrary unitarily invariant norm $\|\cdot\|$.

Let $A \in \{\mathbb{J}, \mathbb{L}\}$. Then recall that $MA \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$ when $A \in \mathbb{S}$. Now we derive the structured backward error of a subspace $\mathcal{X} \subset \mathbb{C}^n$ to be an invariant subspace of a given matrix $A \in \mathbb{S}$. First we consider $A \in \{\mathsf{sym}, \mathsf{skew\text{-}sym}, \mathsf{Herm}, \mathsf{skew\text{-}Herm}\}$.

The structured backward error of $\mathcal{X}$ is defined by

$$\eta^{\mathbb{S}}(\mathcal{X}, A) := \inf_{E \in \mathbb{S}} \{\|E\| : (A + E)\mathcal{X} \subseteq \mathcal{X}\}.$$

For the Frobenius and the spectral norm we denote $\eta^{\mathbb{S}}(\mathcal{X}, A)$ by $\eta_F^{\mathbb{S}}(\mathcal{X}, A)$ and $\eta_2^{\mathbb{S}}(\mathcal{X}, A)$ respectively. We mention that Stewart, [86] obtained $\eta_F^{\mathbb{S}}(\mathcal{X}, A)$ when $A \in \mathsf{Herm}$.

Let $\mathcal{X}$ be a subspace of $\mathbb{C}^n$ and $A \in \mathsf{sym}$. To determine $\eta^{\mathbb{S}}(\mathcal{X}, A)$, we assume that there is an isometry $X \in \mathbb{R}^{n \times k}$ such that $\text{Range}(X) = \mathcal{X}$.

Theorem 2.4.1. *Let* $\mathbb{S} = \mathsf{sym}$ *and* $A \in \mathbb{S}$. *Let* $\mathcal{X}$ *be a subspace of* $\mathbb{C}^n$ *such that there exist an*

TH-789_04612301

isometry $X \in \mathbb{R}^{n\times k}$ *such that* $\mathrm{Range}(X) = \mathcal{X}$. *Then*

$$\eta_F^{\mathbb{S}}(\mathcal{X}, A) = \sqrt{2}\,\|(I - XX^T)AX\|_F, \;\; \eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^T)AX\|_2.$$

Define $E := -XX^TA - AXX^T + 2XX^TAXX^T$. *Then* $E \in \mathbb{S}, \|E\|_{2,F} = \eta_F^{\mathbb{S}}(\mathcal{X}, A)$ *and* $(A + E)\mathcal{X} \subseteq \mathcal{X}$.

Proof: By Theorem 2.1 it is easy to verify that for given $A \in \mathsf{sym}$ and $X \in \mathbb{C}^{n\times k}$ there always exists $E \in \mathsf{sym}$ such that

$$(A + E)X = XD \Rightarrow EX = XD - AX \tag{2.6}$$

for some $D = D^T$ of order k.

Now construct a real isometry Q_1 such that $Q = [X,\ Q_1]$ is orthogonal and $Q_1^T X = 0$. Define $E = Q\begin{bmatrix} E_{11} & E_{12}^T \\ E_{12} & E_{22}\end{bmatrix}Q^T$. Hence from (2.6) we obtain

$$\begin{bmatrix} E_{11} & E_{12}^T \\ E_{12} & E_{22}\end{bmatrix}\begin{bmatrix} X^T \\ Q_1^T\end{bmatrix}X = \begin{bmatrix} X^T \\ Q_1^T\end{bmatrix}(XD - AX)$$

which gives $E_{11} = D - X^TAX,\ E_{12} = -Q_1^TAX$. Therefore we have

$$E = Q\begin{bmatrix} D - X^TAX & (-Q_1^TAX)^T \\ -Q_1^TAX & E_{22}\end{bmatrix}Q^T. \tag{2.7}$$

This gives $\|E\|_F^2 = \|D - X^TAX\|_F^2 + 2\|Q_1^TAX\|_F^2 + \|E_{22}\|_F^2$. To minimize $\|E\|_F$ we fix $E_{22} = 0$.

Consequently we have $\|E\|_F^2 = \|D - X^TAX\|_F^2 + 2\|Q_1^TAX\|_F^2$, where D is a complex symmetric matrix of order k. Setting $D := X^TAX$ we obtain

$$\eta_F(\mathcal{X}, \mathbb{S}) = \sqrt{2}\|Q_1^TAX\|_F = \sqrt{2}\sqrt{\|X^TA\|_F^2 - \|X^TAX\|_F^2}.$$

given by $E = -XX^TA - AXX^T + 2XX^TAXX^T$.

Next, consider the spectral norm. By (2.7) and (2.5) we have

$$\|E\|_2 \geq \left\|\begin{bmatrix} 0 & (-Q_1^TAX)^T \\ -Q_1^TAX & 0\end{bmatrix}\right\|_2 = \|Q_1^TAX\|_2 \tag{2.8}$$

for any $D = D^T$ of order k. Setting $D = X^TAX$, by (2.7) we have

$$E = Q\begin{bmatrix} 0 & (-Q_1^HAX)^T \\ -Q_1^HAX & E_{22}\end{bmatrix}Q^H.$$

Now by DKW Theorem 1.2.5 we construct infinitely many symmetric matrices $E_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$ such that $\|E\|_2 = \|Q_1^TAX\|_2 = \|(I - XX^T)AX\|_2$. Note that the lower bound is attained by this particular choice of D. Then by (2.8) we have

$$\eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^T)AX\|_2.$$

TH-789_04612301

Moreover
$$E_{22} = \mu(I - KK^H)^{1/2} Z (I - \overline{K}K^T)^{1/2}$$
where $K := -\mu^{-1} Q_1^T AX$, $\mu = \|(I - XX^T)AX\|_2$ and $Z = Z^T$ is an arbitrary contraction. In particular, taking $Z = 0$ and simplifying the expression E we obtain the same as that of for Frobenius norm.■

Note that $E \in \mathsf{sym}$ is unique such that $\|E\|_F = \eta_F^{\mathbb{S}}(\mathcal{X}, A)$ and $(A+E)\mathcal{X} \subset \mathcal{X}$. Also notice that for the spectral norm we have

$$E = -XX^T A - AXX^T + 2XX^T AXX^T + \mu Q_1 (I - KK^H)^{1/2} Z (I - \overline{K}K^T)^{1/2} Q_1^T,$$

where $Q_1^T X = 0, K := -\mu^{-1} Q_1^T AX$, $\mu = \|(I - XX^T)AX\|_2$ and $Z = Z^T$ is an arbitrary contraction, such that $\|E\|_2 = \eta_2^{\mathbb{S}}(\mathcal{X}, A)$ and $(A + E)\mathcal{X} \subset \mathcal{X}$.

Corollary 2.4.2. *Let $\mathbb{S} = \mathsf{sym}$. Let $A \in \mathbb{S}$ and $\mathcal{X}$ be a subspace of $\mathbb{C}^n$. Then any $E \in \mathbb{S}$ such that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ is given by*

$$E = XDX^T - XX^T A - AXX^T + XX^T AXX^T + (I - XX^T)Z(I - XX^T)$$

where $D^T = D \in \mathbb{C}^{k\times k}, Z^T = Z \in \mathbb{C}^{(n-k)\times(n-k)}$.

Proof: The proof is followed by simplifying (2.7). ■

Next we consider $\mathcal{X}$ is a subspace of $\mathbb{C}^n$ and $A \in \mathsf{skew\text{-}sym}$. To determine $\eta^{\mathbb{S}}(\mathcal{X}, A)$, we assume that there is an isometry $X \in \mathbb{R}^{n\times k}$ such that $\mathrm{Range}(X) = \mathcal{X}$.

Theorem 2.4.3. *Let $\mathbb{S} = \mathsf{skew\text{-}sym}$ and $A \in \mathbb{S}$. Let $\mathcal{X}$ be a subspace of $\mathbb{C}^n$ such that there exist a real isometry $X \in \mathbb{C}^{n\times k}$ and $\mathrm{Range}(X) = \mathcal{X}$. Then*

$$\eta_F^{\mathbb{S}}(\mathcal{X}, A) = \sqrt{2}\,\|(I - XX^T)AX\|_F, \ \ \eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^T)AX\|_2.$$

Define $E := -XX^T A - AXX^T + 2XX^T AXX^T$. Then $\|E\|_{F,2} = \eta_{2,F}^{\mathbb{S}}(\mathcal{X}, A)$ and $(A+E)\mathcal{X} \subseteq \mathcal{X}$.

Proof: The proof is similar to Theorem 2.4.1. ■

Note that $E \in \mathsf{skew\text{-}sym}$ is a unique matrix such that $\|E\|_F = \eta_F^{\mathbb{S}}(\mathcal{X}, A)$ and $(A+E)\mathcal{X} \subseteq \mathcal{X}$. Also notice that

$$E = XX^T A - AXX^T + 2XX^T AXX^T + \mu Q_1 (I - KK^H)^{1/2} Z (I - \overline{K}K^T)^{1/2} Q_1^T,$$

where $Q_1^T X = 0, K := -\mu^{-1} Q_1^T AX$, $\mu = \|AX\|_2$ and $Z = -Z^T$ is an arbitrary contraction, is such that $E \in \mathsf{skew\text{-}sym}$, $\|E\|_2 = \eta_2^{\mathbb{S}}(\mathcal{X}, A)$ and $(A + E)\mathcal{X} \subseteq \mathcal{X}$.

Corollary 2.4.4. *Let $\mathbb{S} = \mathsf{skew\text{-}sym}$. Let $A \in \mathbb{S}$ and $\mathcal{X}$ be a subspace of $\mathbb{C}^n$. Then any $E \in \mathbb{S}$ such that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ is given by*

$$E = XDX^T + XX^T A + XX^T AXX^T - AXX^T + (I - XX^T)Z(I - XX^T)$$

where $D^T = -D \in \mathbb{C}^{k\times k}, Z^T = -Z \in \mathbb{C}^{(n-k)\times(n-k)}$.

TH-789_04612301

Now we consider $\mathbb{S} = \mathsf{Herm}$.

Theorem 2.4.5. *Let $\mathbb{S} = \mathsf{Herm}$ and $A \in \mathbb{S}$. Let $\mathcal{X}$ be a subspace of A and $X \in \mathbb{C}^{n\times k}$ be an isometry such that* $\text{Range}(X) = \mathcal{X}$. *Then*

$$\eta_F^{\mathbb{S}}(\mathcal{X}, A) = \sqrt{2}\,\|(I - XX^H)AX\|_F, \;\; \eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^H)AX\|_2.$$

Define $E = -XX^HA - AXX^H + 2XX^HAXX^H$. Then $(A+E)\mathcal{X} \subseteq \mathcal{X}$ and $\|E\|_{2,F} = \eta_{2,F}^{\mathbb{S}}(\mathcal{X}, A)$ for both the Frobenius and the spectral norms..

Proof: Let $X := \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix} \in \mathbb{C}^{n\times k}$ be an isometry such that $\text{Range}(X) = \mathcal{X}$. By Theorem 2.1 it follows that there exist $E \in \mathbb{S}$ such that

$$(A+E)X = XD \Rightarrow EX = XD - AX \tag{2.9}$$

for some $D = D^H$ of order k. Now construct a unitary matrix $Q = [X,\, Q_1]$ such that $Q_1^H X = 0$. Define $E = Q\begin{bmatrix} E_{11} & E_{12}^H \\ E_{12} & E_{22} \end{bmatrix} Q^H$. Hence by (2.9) we obtain

$$\begin{bmatrix} E_{11} & E_{12}^H \\ E_{12} & E_{22} \end{bmatrix} \begin{bmatrix} X^H \\ Q_1^T \end{bmatrix} X = \begin{bmatrix} X^H \\ Q_1^H \end{bmatrix} (XD - AX)$$

which gives $E_{11} = D - X^HAX,\; E_{12} = -Q_1^HAX$. Therefore we have

$$E = Q\begin{bmatrix} D - X^HAX & (-Q_1^HAX)^H \\ -Q_1^HAX & E_{22} \end{bmatrix} Q^H. \tag{2.10}$$

This gives $\|E\|_F^2 = \|D - X^HAX\|_F^2 + 2\|Q_1^HAX\|_F^2 + \|E_{22}\|_F^2$. To minimize $\|E\|_F$ we fix $E_{22} = 0$.

Consequently we have $\|E\|_F^2 = \|D - X^HAX\|_F^2 + 2\|Q_1^HAX\|_F^2$, where D is any Hermitian matrix of order k. Setting $D := X^HAX$ we obtain

$$\eta_F^{\mathbb{S}}(\mathcal{X}, A) = \sqrt{2}\|Q_1^HAX\|_F = \sqrt{2}\sqrt{\|X^HA\|_F^2 - \|X^HAX\|_F^2}.$$

Simplifying the expressions of E we obtain $E = -XX^HA - AXX^H + 2XX^HAXX^H$.

Next, consider the spectral norm. By (2.7) and (2.5) we have the following lower bound

$$\|E\|_2 \geq \left\| \begin{bmatrix} 0 & (-Q_1^HAX)^H \\ -Q_1^HAX & 0 \end{bmatrix} \right\|_2 = \|Q_1Q_1^HAX\|_2 = \|(I - XX^H)AX\|_2 \tag{2.11}$$

for any $D = D^H$ of order k. Set $D = X^HAX$. Then by (2.10) we have

$$E = Q\begin{bmatrix} 0 & (-Q_1^HAX)^H \\ -Q_1^HAX & E_{22} \end{bmatrix} Q^H.$$

Using dilation Theorem 1.2.6 for Hermitian matrices we can construct infinitely many Hermitian matrices $E_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$ such that $\|E\|_2 = \|Q_1^HAX\|_2 = \|(I - XX^H)AX\|_2$. Note that the lower bound is attained by this particular choice of D. Consequently, by (2.11) we

TH-789_04612301

obtain

$$\eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^H)AX\|_2.$$

Moreover

$$E_{22} = \mu(I - KK^H)^{1/2} Z (I - KK^H)^{1/2}$$

where $K := -\mu^{-1} Q_1^H AX$, $\mu = \|(I - XX^H)AX\|_2$ and $Z = Z^H$ is an arbitrary contraction. In particular, taking $Z = 0$ and simplifying the expression E we obtain the same as that of for Frobenius norm.■

It follows from the proof that E is unique such that $\|E\_F = \eta_F^{\mathbb{S}}(\mathcal{X}, A)$ and $(A+E)\mathcal{X} \subseteq \mathcal{X}$. For the spectral norm, E is given by

$$E = -XX^H A - AXX^H + 2XX^H AXX^H + \mu Q_1 (I - KK^H)^{1/2} Z (I - KK^H)^{1/2} Q_1^H,$$

where $Q_1^H X = 0, K := -\mu^{-1} Q_1^H AX$, $\mu = \|(I - XX^H)AX\|_2$ and $Z = Z^H$ is an arbitrary contraction, such that $\|E\|_2 = \eta_2^{\mathbb{S}}(\mathcal{X}, A)$ and $(A + E)\mathcal{X} \subseteq \mathcal{X}$.

Corollary 2.4.6. *Let $\mathbb{S} = \mathsf{Herm}$ and $A \in \mathbb{S}$. Let $\mathcal{X}$ be a subspace of $\mathbb{C}^n$ and $X \in \mathbb{C}^{n\times k}$ is an isometry such that* Range$(X) = \mathcal{X}$. *Then any $E \in \mathbb{S}$ such that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ is given by*

$$E = XDX^H - XX^H A - AXX^H + XX^H AXX^H + (I - XX^H)Z(I - XX^H)$$

where $D^H = D \in \mathbb{C}^{k\times k}, Z^H = Z \in \mathbb{C}^{(n-k)\times(n-k)}$.

Proof: The proof is followed by simplifying (2.10). ■

Now we consider $\mathbb{S} = \mathsf{skew\text{-}Herm}$.

Theorem 2.4.7. *Let $\mathbb{S} = \mathsf{skew\text{-}Herm}$. $\mathcal{X}$ be a subspace of $\mathbb{C}^n$ and $X \in \mathbb{C}^{n\times k}$ is an isometry such that* Range$(X) = \mathcal{X}$. *Then*

$$\eta_F^{\mathbb{S}}(\mathcal{X}, A) = \sqrt{2}\,\|(I - XX^H)AX\|_F, \;\; \eta_2^{\mathbb{S}}(\mathcal{X}, A) = \|(I - XX^H)AX\|_2.$$

Define $E = XX^H A - AXX^H$. Then $E \in \mathbb{S}$ such that $(A+E)\mathcal{X} \subseteq \mathcal{X}$ and $\|E\|_{2,F} = \eta_{2,F}^{\mathbb{S}}(\mathcal{X}, A)$ for both the Frobenius and the spectral norms.

Proof: The proof is followed by Theorem 2.4.5 ■

Let $\mathbb{S} = \mathsf{skew\text{-}Herm}$. Define $E = XX^H A - AXX^H$. Then $E \in \mathbb{S}$ is unique such that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ and $\|E\|_F = \eta_F^{\mathbb{S}}(\mathcal{X}, A)$. Next define

$$E = XX^H A - AXX^H + \mu Q_1 (I - KK^H)^{1/2} Z (I - KK^H)^{1/2} Q_1^H,$$

where $Q_1^H X = 0, K := -\mu^{-1} Q_1^H AX$, $\mu = \|(I - XX^H)AX\|_2$ and $Z = -Z^H$ is an arbitrary contraction. then notice that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ and $\|E\|_2 = \eta_2^{\mathbb{S}}(\mathcal{X}, A)$.

Corollary 2.4.8. *Let $\mathbb{S} = \mathsf{skew\text{-}Herm}$. Let $A \in \mathbb{S}$ and $\mathcal{X}$ be a subspace of $\mathbb{C}^n$. Then any $E \in \mathbb{S}$ such that $(A + E)\mathcal{X} \subseteq \mathcal{X}$ is given by*

$$E = XDX^H + XX^H A - AXX^H - XX^H AXX^H + (I - XX^H)Z(I - XX^H)$$

TH-789_04612301

where $D^H = -D \in \mathbb{C}^{k\times k}, Z^H = -Z \in \mathbb{C}^{(n-k)\times(n-k)}$.

We mention that the results obtained above can easily be extended to the case when $\mathbb{S} \subset \{\mathbb{J}, \mathbb{L}\}$.

2.5 Structured inverse least square problem

Note that $\mathbb{S}(X, B)$ could be an empty set whenever the pair $(X, B) \in \mathbb{C}^{n\times k} \times \mathbb{C}^{n\times k}$ does not satisfy the necessary and sufficient conditions for the existence of the solution of structured mapping problem for matrices, given in Theorem 2.2.1. Such a pair gives rise to the following structured inverse least-squared problem (SILSP).

Problem. Determine $\alpha_{\mathbb{S}} = \min\{\|AX - B\|_F : A \in \mathbb{S}\}$ and all $A \in \mathbb{S}$ such that $\|A\|_F = \min\{\|ZX - B\|_F : Z \in \mathbb{S}\}$.

A lot of interest have been paid to resolve this problem, see for example in [107], Zhang et al. solved this problem for anti-Hermitian generalized Hamiltonian matrices. We solve the problem for $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$.

We follow the same routine to solve SILSP for $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$ as we did for solving the structured mapping problem by considering the four particular structures namely sym, skew-sym, Herm, skew-Herm first. To begin with we consider SILSP when $X = x \in \mathbb{K}^n$ and $B = b \in \mathbb{K}^n$.

For any $x, b \in \mathbb{K}^n \setminus \{0\}$ we always have $A \in \text{sym} \subset \mathbb{K}^{n\times n}$ such that $Ax = b$. Therefore, for $\mathbb{S}$, the space of symmetric matrices we obtain $\alpha_{\mathbb{S}} = 0$. Next we consider $\mathbb{S} = \text{skew-sym}$.

Theorem 2.5.1. *Let* $\mathbb{S} = \text{skew-sym}$. *Then for any* $x, b \in \mathbb{K}^n$ *we have*

$$\alpha_{\mathbb{S}} = \frac{|x^T b|}{\|x\|}.$$

Proof: Assume that $x, b \in \mathbb{K}^n$. Construct a unitary matrix $Q = [x/\|x\| \; Q_1] \in \mathbb{K}^{n\times n}$ such that $Q_1^H x = 0$. Then construct the skew-symmetric matrix $A = \overline{Q}\begin{bmatrix} 0 & -a_1^T \\ a_1 & A_1 \end{bmatrix} Q^H$. Then we obtain

$$\min_{A\in\mathbb{S}} \|Ax - b\|_F^2 = \min_{A\in\mathbb{S}} \left\| \begin{bmatrix} x^T b/\|x\| \\ a_1\|x\| - Q_1^T b \end{bmatrix} \right\|_F^2 = |x^T b|^2/\|x\|^2 + \min_{a_1\in\mathbb{K}^{n-1}} \|a_1\|x\| - Q_1^T b\|_F^2.$$

Choose $a_1 = Q_1^T b/\|x\|$. Consequently, we obtain $\alpha_{\mathbb{S}} = |x^T b|/\|x\|$. Next we have,

$$\begin{aligned} A &= \overline{Q} \begin{bmatrix} 0 & -(Q_1^T b/\|x\|)^T \\ Q_1^T b/\|x\| & A_1 \end{bmatrix} Q^H \\ &= \frac{1}{\|x\|}[bx^H - \overline{x} b^T] + (I - xx^H)^T Z (I - xx^H) \end{aligned}$$

where $Z = -Z^T \in \mathbb{C}^{n\times n}$. ■

Now we consider Hermitian matrices.

TH-789_04612301

Theorem 2.5.2. *Let* $\mathbb{S} = \mathsf{Herm}$. *Then for any* $x, b \in \mathbb{C}^n$ *we have*

$$\alpha_{\mathbb{S}} = \frac{|\mathsf{im}(x^H b)|}{\|x\|}.$$

Proof: Assume that $x, b \in \mathbb{C}^n$. Construct a unitary matrix $Q = [x/\|x\| \; Q_1] \in \mathbb{C}^{n\times n}$ such that $Q_1^H x = 0$. Then construct the Hermitian matrix $A = Q \begin{bmatrix} a_{11} & a_1^H \\ a_1 & A_1 \end{bmatrix} Q^H$. Then we obtain

$$\min_{A\in\mathbb{S}} \|Ax - b\|_F^2 = \min_{A\in\mathbb{S}} \left\| \begin{bmatrix} a_{11}\|x\| - x^H b/\|x\| \\ a_1\|x\| - Q_1^H b \end{bmatrix} \right\|_F^2 = |a_{11}\|x\| - x^H b/\|x\||^2 + \min_{a_1\in\mathbb{K}^{n-1}} \|a_1\|x\| - Q_1^H b\|_F^2.$$

Choose $a_1 = Q_1^H b/\|x\|$. Next $\min_{a_{11}\in\mathbb{R}} |a_{11}\|x\| - x^H b/\|x\||^2$ is minimized when $a_{11} = \frac{\mathsf{re}(x^H b)}{\|x\|^2}$. Consequently, we obtain $\alpha_{\mathbb{S}} = |\mathsf{im}(x^H b)|/\|x\|$. Next we have,

$$\begin{aligned} A &= Q \begin{bmatrix} \frac{\mathsf{re}(x^H b)}{\|x\|^2} & (Q_1^H b/\|x\|)^H \\ Q_1^H b/\|x\| & A_1 \end{bmatrix} Q^H \\ &= \frac{\mathsf{re}(x^H b)}{\|x\|^4} x x^H + \frac{1}{\|x\|^2}[x b^H (I - x x^H) + (I - x x^H) b x^H] + (I - x x^H) Z (I - x x^H) \end{aligned}$$

where $Z = Z^H \in \mathbb{C}^{n\times n}$. ■

In a similar fashion we can obtain the solution of SILSP for skew-Hermitian matrices as follows.

Theorem 2.5.3. *Let* $\mathbb{S} = \mathsf{skew\text{-}Herm}$. *Then for any* $x, b \in \mathbb{C}^n$ *we have*

$$\alpha_{\mathbb{S}} = \frac{|\mathsf{re}(x^H b)|}{\|x\|}.$$

Proof: The proof is similar to Theorem 2.5.2.■

Now assume that $\mathbb{S} = \mathbb{J}$ or $\mathbb{S} = \mathbb{L}$ where $\mathbb{J}$ and $\mathbb{L}$ are the Jordan algebra and Lie algebra corresponding to orthosymmetric bilinear / sesquilinear scalar product $\langle\cdot,\cdot\rangle_M$. Then its evident by Theorem 2.5.1, Theorem 2.5.2 and Theorem 2.5.3 thet whenever $\mathbb{S}(x, b) = \emptyset$ for a given $x, b \in \mathbb{C}^n \setminus \{0\}$ then

$$\min_{A\in\mathbb{S}} \|Ax - b\| = \begin{cases} \dfrac{|\langle x, b\rangle_M|}{\|x\|_2}, & \text{if} \quad (MA)^T = -MA \\ \dfrac{|\mathsf{im}\langle x, b\rangle_M|}{\|x\|_2}, & \text{if} \quad (MA)^H = MA \\ \dfrac{|\mathsf{re}\langle x, b\rangle_M|}{\|x\|_2}, & \text{if} \quad (MA)^H = -MA. \end{cases}$$

Now we consider SILSP for matrices. Before that we prove the following result which will be used in the subsequent development.

Lemma 2.5.4. *Let* $\alpha, \beta > 0$ *and* $b_1, b_2 \in \mathbb{C}$. *Then* $\min_{x\in\mathbb{C}}(|x\alpha - b_1|^2 + |x\beta - b_2|^2)$ *is given by* $x = \dfrac{\alpha b_1 + \beta b_2}{\alpha^2 + \beta^2}$.

TH-789_04612301

Proof: Assume that $x = x_1 + ix_2$, $b_1 = b_{11} + ib_{12}$, $b_2 = b_{21} + ib_{22} \in \mathbb{C}$. Then define

$$\begin{aligned}\phi(x_1, x_2) &= |x\alpha - b_1|^2 + |x\beta - b_2|^2 \\ &= (\alpha^2 + \beta^2)(x_1^2 + x_2^2) + (b_{11}^2 + b_{12}^2) + (b_{21}^2 + b_{22}^2) - 2\alpha(x_1 b_{11} + x_2 b_{12}) - 2\beta(x_1 b_{21} + x_2 b_{22}).\end{aligned}$$

Setting $\dfrac{\partial \phi(x_1, x_2)}{\partial x_i} = 0, \ i = 1, 2$ we obtain the stationary points

$$x_1 = \frac{\alpha b_{11} + \beta b_{21}}{\alpha^2 + \beta^2},\ x_2 = \frac{\alpha b_{12} + \beta b_{22}}{\alpha^2 + \beta^2}.$$

This gives the relative minimum as the Hessian matrix

$$\begin{pmatrix} \dfrac{\partial^2 \phi}{\partial x_1^2} & \dfrac{\partial^2 \phi}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 \phi}{\partial x_2 \partial x_1} & \dfrac{\partial^2 \phi}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 2(\alpha^2 + \beta^2) & 0 \\ 0 & 2(\alpha^2 + \beta^2) \end{pmatrix}$$

is positive definite. Consequently we obtain the desired result.■

Theorem 2.5.5. *Let $\mathbb{S}$ be the space of symmetric matrices and $X, B \in \mathbb{K}^{n\times k}$ with $\operatorname{rank}(X) = r$. Assume that the SVD of $X = U\Sigma V^H$ where $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}, \Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r), \sigma_1 > \ldots > \sigma_r > 0, U = [U_1\ U_2], V = [V_1\ V_2]$. Then*

$$\min_{A\in\mathbb{S}} \|AX - B\|_F = \|K \circ (U_1^T B V_1 \Sigma_1 + \Sigma_1 V_1^T B^T U_1)\Sigma_1 - U_1^T B V_1\|_F^2 + \|U_1^T B V_2\|_F^2 + \|U_2^T B V_2\|_F^2$$

is attained at

$$A = \overline{U}_1[K \circ U_1^T B V_1 \Sigma_1 + K \circ \Sigma_1 V_1^T B^T U_1]U_1^H + (BX^\dagger)^T (I - XX^\dagger) + (I - XX^\dagger)^T BX^\dagger + (I - XX^\dagger)^T Z (I - XX^\dagger)$$

where $Z^T = Z \in \mathbb{C}^{r\times r}, K = [k_{ij}], k_{ij} = \frac{1}{\sigma_i^2 + \sigma_j^2}$ and $\circ$ denotes the Hadamard product.

Proof: Assume that $X, K \in \mathbb{K}^{n\times k}$ with $\operatorname{rank}(X) = r$. Consider the SVD $X = U\Sigma V^H$. Define a symmetric linear map $A : \operatorname{Range}(X) \oplus \operatorname{Range}(X)^\perp \to \operatorname{Range}(\overline{X}) \oplus \operatorname{Range}(\overline{X})^\perp$. From the SVD of X given above it is easily seen that (U_1, U_2) and $(\overline{U}_1, \overline{U}_2)$ are bases of $\operatorname{Range}(X) \oplus \operatorname{Range}(X)^\perp$ and $\operatorname{Range}(\overline{X}) \oplus \operatorname{Range}(\overline{X})^\perp$ respectively. The block matrix representation of A is of the form

$$A = \begin{bmatrix} \overline{U}_1 & \overline{U}_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix}$$

where A_{11}, A_{12}, A_{22} are of compatible sizes. Consequently we obtain

$$\begin{aligned}\|AX - B\|_F^2 &= \|U^T A U U^H X - U^T B\|_F^2 \\ &= \left\| \begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} U_1^H X \\ 0 \end{bmatrix} - \begin{bmatrix} U_1^T B \\ U_2^T B \end{bmatrix} \right\|_F^2 \\ &= \|A_{11} U_1^H X - U_1^T B\|_F^2 + \|A_{12} U_1^H X - U_2^T B\|_F^2.\end{aligned}$$

TH-789_04612301

Then

$$\begin{aligned}\min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HX-U_2^TB\|_F^2 &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HU\Sigma V^H-U_2^TB\|_F^2\\ &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HU\Sigma-U_2^TBV\|_F^2\\ &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}\Sigma_1-U_2^TBV_1\|_F^2+\|U_2^TBV_2\|_F^2.\end{aligned}$$

Moreover $\|A_{12}\Sigma_1-U_2^TBV_1\|_F^2$ is minimized if and only if $A_{12}\Sigma_1-U_2^TBV_1=0$ i.e. if and only if $A_{12}=U_2^TBV_1\Sigma_1^{-1}$. Similarly we obtain

$$\min_{A_{11}\in\mathbb{C}^{r\times r},A_{11}^T=A_{11}}\|A_{11}U_1^HX-U_1^TB\|_F^2=\min_{A_{11}\in\mathbb{C}^{r\times r},A_{11}^T=A_{11}}\|A_{11}\Sigma_1-U_1^TBV_1\|_F^2+\|U_1^TBV_2\|_F^2.$$

Further assume that $A_{11}=[a_{ij}]$, $U_1^TBV_1=[b_{ij}]$. Consequently we have,

$$\begin{aligned}\|A_{11}\Sigma_1-U_1^TBV_1\|_F^2 &= \sum_{j\leq i,j=1}^{r}\sum_{i=1}^{r}(|a_{ij}\sigma_i-b_{ij}|^2+|a_{ji}\sigma_j-b_{ji}|^2)\\ &= \sum_{j\leq i,j=1}^{r}\sum_{i=1}^{r}(|a_{ij}\sigma_i-b_{ij}|^2+|a_{ij}\sigma_j-b_{ji}|^2)\end{aligned}$$

The desired minimum can be obtained by minimizing $|a_{ij}\sigma_i-b_{ij}|^2+|a_{ij}\sigma_j-b_{ji}|^2$, for all $i,j=1:r$. By Lemma 2.5.4 the minimum can be obtained by

$$a_{ij}=\frac{\sigma_ib_{ij}+b_{ji}\sigma_j}{\sigma_i^2+\sigma_j^2},\ a_{ij}=a_{ji},\forall i,j=1:r.$$

Hence $\|A_{11}\Sigma_1-U_1^TBV_1\|_F^2$ can be minimized by

$$A_{11}=K\circ(U_1^TBV_1\Sigma_1+\Sigma_1V_1^TB^TU_1)$$

where $K=[k_{ij}],k_{ij}=\frac{1}{\sigma_i^2+\sigma_j^2}$ and $\circ$ denotes the Hadamard product. Consequently, we obtain

$$\begin{aligned}A &= \overline{U}\begin{bmatrix}K\circ(U_1^TBV_1\Sigma_1+\Sigma_1V_1^TB^TU_1) & \Sigma_1^{-1}V_1^TB^TU_2\\ U_2^TBV_1\Sigma_1^{-1} & A_{22}\end{bmatrix}U^H\\ &= \overline{U}_1[K\circ(U_1^TBV_1\Sigma_1+\Sigma_1V_1^TB^TU_1)]U_1^H+\overline{U}_1\Sigma_1^{-1}V_1^{-1}B^TU_2U_2^H+\overline{U}_1U_2^TBV_1\Sigma_1^{-1}U_1^H+\overline{U}_2A_{22}U_2^H\\ &= \overline{U}_1[K\circ(U_1^TBV_1\Sigma_1+\Sigma_1V_1^TB^TU_1)]U_1^H+(BX^\dagger)^T(I-XX^\dagger)+(I-XX^\dagger)^TBX^\dagger\\ &\quad +(I-XX^\dagger)^TZ(I-XX^\dagger),\ Z^T=Z\in\mathbb{C}^{r\times r}\end{aligned}$$

which gives the desired result.■

Now we consider skew-symmetric matrices.

Theorem 2.5.6. *Let $\mathbb{S}$ be the space of skew-symmetric matrices and $X,B\in\mathbb{K}^{n\times k}$ with* $\mathrm{rank}(X)=r$. *Assume that the SVD of $X=U\Sigma V^H$ where $\Sigma=\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}$*, $\Sigma_1=\mathrm{diag}(\sigma_1,\ldots,\sigma_r),\sigma_1>\ldots>\sigma_r>0,U=[U_1\ U_2],V=[V_1\ V_2]$. *Then*

$$\alpha_{\mathbb{S}}=\|K\circ(U_1^TBV_1\Sigma_1-\Sigma_1V_1^TB^TU_1)\Sigma_1-U_1^TBV_1\|_F^2+\|U_1^TBV_2\|_F^2+\|U_2^TBV_2\|_F^2$$

TH-789_04612301

is attained at

$$A=\overline{U}_1[K\circ U_1^TBV_1\Sigma_1-K\circ\Sigma_1V_1^TB^TU_1]U_1^H-(BX^\dagger)^T(I-XX^\dagger)+(I-XX^\dagger)^TBX^\dagger+(I-XX^\dagger)^TZ(I-XX^\dagger)$$

where $Z^T=-Z\in\mathbb{C}^{r\times r}K=[k_{ij}],k_{ij}=\frac{1}{\sigma_i^2+\sigma_j^2}$ *and* $\circ$ *denotes the Hadamard product.*

Proof: The proof is similar to the proof for symmetric case.■

Next, we consider Hermitian matrices.

Theorem 2.5.7. *Let* $\mathbb{S}$ *be the space of Hermitian matrices and* $X,B\in\mathbb{K}^{n\times k}$ *with* $\mathrm{rank}(X)=r$. *Assume that the SVD of* $X=U\Sigma V^H$ *where* $\Sigma=\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix},\Sigma_1=\mathrm{diag}(\sigma_1,\ldots,\sigma_r),\sigma_1>\ldots>\sigma_r>0,U=[U_1\ U_2],V=[V_1\ V_2]$. *Then*

$$\alpha_{\mathbb{S}}=\|K\circ(U_1^HBV_1\Sigma_1+\Sigma_1V_1^HB^HU_1)\Sigma_1-U_1^HBV_1\|_F^2+\|U_1^HBV_2\|_F^2+\|U_2^HBV_2\|_F^2$$

is attained at

$$A=U_1[K\circ U_1^HBV_1\Sigma_1+K\circ\Sigma_1V_1^HB^HU_1]U_1^H+(BX^\dagger)^H(I-XX^\dagger)+(I-XX^\dagger)BX^\dagger+(I-XX^\dagger)Z(I-XX^\dagger)$$

where $Z^H=Z\in\mathbb{C}^{r\times r},K=[k_{ij}],k_{ij}=\frac{1}{\sigma_i^2+\sigma_j^2}$ *and* $\circ$ *denotes the Hadamard product.*

Proof: Assume that $X,K\in\mathbb{K}^{n\times k}$ with $\mathrm{rank}(X)=r$. Consider the SVD $X=U\Sigma V^H$. Define a Hermitian linear map $A:\mathrm{Range}(X)\oplus\mathrm{Range}(X)^\perp\to\mathrm{Range}(X)\oplus\mathrm{Range}(X)^\perp$. From the SVD of X given above it is easily seen that (U_1,U_2) and (U_1,U_2) are bases of $\mathrm{Range}(X)\oplus\mathrm{Range}(X)^\perp$ and $\mathrm{Range}(X)\oplus\mathrm{Range}(X)^\perp$ respectively. The block matrix representation of A is of the form

$$A=U\begin{bmatrix}A_{11} & A_{12}^H\\ A_{12} & A_{22}\end{bmatrix}U^H.$$

Consequently we obtain

$$\begin{aligned}\|AX-B\|_F^2 &= \|U^HAUU^HX-U^HB\|_F^2\\ &= \left\|\begin{bmatrix}A_{11} & A_{12}^H\\ A_{12} & A_{22}\end{bmatrix}\begin{bmatrix}U_1^HX\\ 0\end{bmatrix}-\begin{bmatrix}U_1^HB\\ U_2^HB\end{bmatrix}\right\|_F^2\\ &= \|A_{11}U_1^HX-U_1^HB\|_F^2+\|A_{12}U_1^HX-U_2^HB\|_F^2.\end{aligned}$$

Then,

$$\begin{aligned}\min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HX-U_2^HB\|_F^2 &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HU\Sigma V^H-U_2^HB\|_F^2\\ &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}U_1^HU\Sigma-U_2^HBV\|_F^2\\ &= \min_{A_{12}\in\mathbb{C}^{(n-r)\times r}}\|A_{12}\Sigma_1-U_2^HBV_1\|_F^2+\|U_2^HBV_2\|_F^2.\end{aligned}$$

Now $\|A_{12}\Sigma_1-U_2^HBV_1\|_F^2$ is minimized if and only if $A_{12}\Sigma_1-U_2^HBV_1=0$ i.e. if and only if $A_{12}=U_2^HBV_1\Sigma_1^{-1}$. Similarly we have,

$$\min_{A_{11}\in\mathbb{C}^{r\times r},A_{11}^H=A_{11}},\|A_{11}U_1^HX-U_1^HB\|_F^2=\min_{A_{11}\in\mathbb{C}^{r\times r},A_{11}^H=A_{11}}\|A_{11}\Sigma_1-U_1^HBV_1\|_F^2+\|U_1^HBV_2\|_F^2.$$

TH-789_04612301

Further let $A = [a_{ij}], U_1^H BV_1 = [b_{ij}] \in \mathbb{C}^{r\times r}$. Consequently we have

$$\begin{aligned}\|A_{11}\Sigma_1 - U_1^H BV_1\|_F^2 &= \sum_{j\le i,j=1}^{r}\sum_{i=1}^{r}(|a_{ij}\sigma_i - b_{ij}|^2 + |\bar{a}_{ij}\sigma_j - b_{ji}|^2)\\ &= \sum_{j\le i,j=1}^{r}\sum_{i=1}^{r}(|a_{ij}\sigma_i - b_{ij}|^2 + |a_{ij}\sigma_j - \bar{b}_{ji}|^2).\end{aligned}$$

Now the desired minimum can be obtained by minimizing $|a_{ij}\sigma_i - b_{ij}|^2 + |a_{ij}\sigma_j - \bar{b}_{ji}|^2$ for all $i, j = 1 : r$. By Lemma 2.5.4 the minimum can be obtained by

$$a_{ij} = \frac{\sigma_i b_{ij} + \sigma_j \bar{b}_{ji}}{\sigma_i^2 + \sigma_j^2},\ a_{ji} = \bar{a}_{ij},\ \forall i, j = 1 : r.$$

Hence we obtain that $\|A_{11}\Sigma_1 - U_1^H BV_1\|_F^2$ is minimized by

$$A_{11} = K \circ (U_1^H BV_1\Sigma_1 + \Sigma_1 V_1^H B^H U_1)$$

where $K = [k_{ij}], k_{ij} = \frac{1}{\sigma_i^2+\sigma_j^2}$ and $\circ$ denotes the Hadamard product. Consequently, we obtain

$$\begin{aligned}A &= U\begin{bmatrix} K\circ(U_1^H BV_1\Sigma_1 - \Sigma_1 V_1^H B^H U_1) & +\Sigma_1^{-1}V_1^H B^H U_2\\ U_2^H BV_1\Sigma_1^{-1} & A_{22}\end{bmatrix}U^H\\ &= U_1[K\circ(U_1^H BV_1\Sigma_1 + \Sigma_1 V_1^H B^H U_1)]U_1^H + U_1\Sigma_1^{-1}V_1^{-1}B^H U_2 U_2^H + U_1 U_2^H BV_1\Sigma_1^{-1}U_1^H + U_2 A_{22} U_2^H\\ &= U_1[K\circ(U_1^H BV_1\Sigma_1 + \Sigma_1 V_1^H B^H U_1)]U_1^H + (BX^\dagger)^H(I - XX^\dagger) + (I - XX^\dagger)BX^\dagger\\ &\quad +(I - XX^\dagger)^H Z(I - XX^\dagger),\ Z^H = Z \in \mathbb{C}^{r\times r}\end{aligned}$$

which gives the desired result.■

Now consider skew-Hermitian matrices.

Theorem 2.5.8. *Let $\mathbb{S}$ be the space of skew-Hermitian matrices and $X, B \in \mathbb{K}^{n\times k}$ with* $\text{rank}(X) = r$. *Assume that the SVD of $X = U\Sigma V^H$ where $\Sigma = \begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}, \Sigma_1 = \text{diag}(\sigma_1, \ldots, \sigma_r), \sigma_1 > \ldots > \sigma_r > 0, U = [U_1\ U_2], V = [V_1\ V_2]$. Then*

$$\alpha_{\mathbb{S}} = \|K\circ(U_1^H BV_1\Sigma_1 - \Sigma_1 V_1^H B^H U_1)\Sigma_1 - U_1^H BV_1\|_F^2 + \|U_1^H BV_2\|_F^2 + \|U_2^H BV_2\|_F^2$$

is attained at

$$A = U_1[K\circ U_1^H BV_1\Sigma_1 - K\circ\Sigma_1 V_1^H B^H U_1]U_1^H - (BX^\dagger)^H(I-XX^\dagger) + (I-XX^\dagger)BX^\dagger + (I-XX^\dagger)Z(I-XX^\dagger)$$

where $Z^H = -Z \in \mathbb{C}^{r\times r}, K = [k_{ij}], k_{ij} = \frac{1}{\sigma_i^2+\sigma_j^2}$ and $\circ$ denotes the Hadamard product.

Proof: The proof is similar to the proof for Hermitian case.■

Now consider $\mathbb{S} \in \{\mathbb{J}, \mathbb{L}\}$. Then for any given $X, B \in \mathbb{C}^{n\times k}$ we have $\|AX - B\|_F = \|MAX - MB\|_F$. Therefore the SILSP problem can be resolved for $\mathbb{S}$ just replacing B by MB. Further the matrix $A \in \mathbb{S}$ which produces the minimum, can be obtained by redefining it as MA.

TH-789_04612301

Chapter 3

Structured backward errors and pseudospectra of structured matrix pencils

Structured backward perturbation analysis plays an important role in the accuracy assessment of computed eigenelements of structured eigenvalue problems. We undertake a detailed structured backward perturbation analysis of approximate eigenelements of linearly structured matrix pencils. The structures we consider include, for example, symmetric, skew-symmetric, Hermitian, skew-Hermitian, even, odd, palindromic and Hamiltonian matrix pencils. We also analyze structured backward errors of approximate eigenvalues and structured pseudospectra of structured matrix pencils.

3.1 Introduction

Backward perturbation analysis determines the smallest perturbation for which a computed solution is an exact solution of the perturbed problem. On the other hand, condition numbers measure the sensitivity of solutions to small perturbations in the data of the problem. Thus, backward errors when combined with condition numbers provide an approximate upper bounds on the errors in the computed solutions.

With a view to preserving structures and their associated properties, structured preserving algorithms for structured eigenproblems have been proposed in the literature (see, for example, [9, 10, 18, 46, 74, 75] and the references therein). Consequently, there is a growing interest in the structured perturbation analysis of structured eigenproblems (see, for example, [16, 38, 51, 54, 81, 95] for sensitivity analysis of structured eigenproblems).

The main purpose of this chapter is to undertake a detailed structured backward perturbation analysis of approximate eigenelements of linearly structured matrix pencils. Needless to mention that structured backward errors when combined with structured condition numbers provide an approximate upper bounds on the errors in the computed eigenelements. Hence structured backward perturbation analysis plays an important role in the accuracy assessment of approximate eigenelements of structured pencils. Further, it also plays an important role in the selection of an optimum structured linearization of a structured matrix polynomial.

TH-789_04612301

This assumes significance due to the fact that linearization is a standard approach to solving a polynomial eigenvalue problem (see, for example, [39, 41] and the references therein).

We consider regular matrix pencils of the form $\mathrm{L}(\lambda) = A + \lambda B$, where A and B are square matrices of size n. We assume L to be linearly structured, that is, L to be an element of a real or a complex linear subspace $\mathbb{S}$ of the space of pencils. More specifically, we consider ten special classes of linearly structured pencils, namely, T-symmetric, T-skew-symmetric, T-odd, T-even, T-palindromic, H-Hermitian, H-skew-Hermitian, H-even and H-odd and H-palindromic. These structures, defined in the next section, are prototypes of structured pencils which occur in many applications (see, [40, 75] and the references therein). We also consider $\mathbb{S}$ to be the space of pencils whose coefficient matrices are elements of Jordan and/or Lie algebras associated with the scalar product $(x, y) \mapsto y^T M x$ or $(x, y) \mapsto y^H M x$, where M is unitary and $M^T = \pm M$ or $M^H = \pm M$. For example, when $M := \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, the Lie and Jordan algebras associated with the scalar product $(x, y) \mapsto y^H M x$ consist of Hamiltonian and skew-Hamiltonian matrices, respectively. The structures so considered encompass a wide variety of structured pencils and, in particular, includes pencils whose coefficient matrices are Hamiltonian and skew-Hamiltonian. We show, however, that analyzing these wide classes of structured pencils ultimately boils down to analyzing one of the ten special classes of structured pencils considered above. Consequently, we consider these ten special classes of structured pencils and investigate structured backward perturbation analysis of approximate eigenelements.

So, let $\mathbb{S}$ be the space of pencils having one of the ten structures. Let $\mathrm{L} \in \mathbb{S}$ and $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ with $x^H x = 1$. Then we define the structured backward error $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ of (λ, x) by

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) := \inf\{|||\triangle\mathrm{L}||| : \triangle\mathrm{L} \in \mathbb{S} \text{ and } \mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0\}.$$

Here the pencil norm $|||\mathrm{L}|||$ is given by $|||\mathrm{L}||| := \sqrt{\|A\|^2 + \|B\|^2}$, where $\mathrm{L}(z) = A + zB$ and $\|\cdot\|$ is either the spectral norm or the Frobenius norm on $\mathbb{C}^{n\times n}$. The main contributions of this chapter are as follows.

Given $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ with $x^H x = 1$ and $\mathrm{L} \in \mathbb{S}$, we show that there is a pencil $\mathrm{K} \in \mathbb{S}$ such that $\mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0$. Consequently, $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) < \infty$. We determine $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ and construct a pencil $\triangle\mathrm{L} \in \mathbb{S}$ such that $|||\triangle\mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ and $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$. Moreover, we show that $\triangle\mathrm{L}$ is unique for the Frobenius norm on $\mathbb{C}^{n\times n}$ but there are infinitely many such $\triangle\mathrm{L}$ for the spectral norm on $\mathbb{C}^{n\times n}$. Further, for the spectral norm, we show how to construct all such $\triangle\mathrm{L}$. In either case, we show that if $\mathrm{K} \in \mathbb{S}$ is such that $\mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K} = \triangle\mathrm{L} + (I - xx^H)^* \, \mathrm{N}(I - xx^H)$ for some $\mathrm{N} \in \mathbb{S}$, where $(I - xx^H)^*$ denotes the transpose or the conjugate transpose of $(I - xx^H)$ depending upon the structure defined by $\mathbb{S}$. Furthermore, we show that the unstructured backward error $\eta(\lambda, x, \mathrm{L})$ of (λ, x) is a lower bound of $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ and is attained by $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ for certain $\lambda \in \mathbb{C}$. However, $\eta(\lambda, x, \mathrm{L}) \neq \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ for most $\lambda \in \mathbb{C}$.

Next, we consider structured pseudospectra of structured matrix pencils. It is a well known fact that pseudospectra of matrices and matrix pencils are powerful tools for sensitivity and perturbation analysis (see, [100] and the references therein). We consider structured and

TH-789_04612301

unstructured ϵ-pseudospectra

$$\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) := \{\lambda \in \mathbb{C} : \eta^{\mathbb{S}}(\lambda, \mathrm{L}) \leq \epsilon\} \text{ and } \sigma_\epsilon(\mathrm{L}) := \{\lambda \in \mathbb{C} : \eta(\lambda, \mathrm{L}) \leq \epsilon\}$$

of L, where $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) := \min_{x^H x=1} \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ and $\eta(\lambda, \mathrm{L}) := \min_{x^H x=1} \eta(\lambda, x, \mathrm{L})$, respectively, are structured and unstructured backward errors of an approximate eigenvalue λ. When L is T-symmetric or T-skew-symmetric pencils, we show that $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \eta(\lambda, \mathrm{L})$ for the spectral norm and $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \sqrt{2}\,\eta(\lambda, \mathrm{L})$ for the Frobenius norm. Consequently, for these structures, we show that $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \sigma_\epsilon(\mathrm{L})$ for the spectral norm and $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \sigma_{\epsilon/\sqrt{2}}(\mathrm{L})$ for the Frobenius norm. For the rest of the structures, we show that there is a set $\Omega \subset \mathbb{C}$ such that $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap \Omega = \sigma_\epsilon(\mathrm{L}) \cap \Omega$. For example, $\Omega = \mathbb{R}$ when L is H-Hermitian or H-skew-Hermitian and $\Omega = i\mathbb{R}$ when L is H-even or H-odd. Often the spectrum of L is symmetric with respect to Ω. When Ω does not contain an eigenvalue of L, it is of practical importance to determine the smallest perturbation $\triangle\mathrm{L} \in \mathbb{S}$ of L such that $\mathrm{L} + \triangle\mathrm{L}$ has an eigenvalue in Ω. We show how to construct such a $\triangle\mathrm{L}$. Indeed, we show that the equality $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap \Omega = \sigma_\epsilon(\mathrm{L}) \cap \Omega$ plays a crucial role in the construction of such a $\triangle\mathrm{L}$.

3.2 Structured matrix pencils

We consider n-by-n matrix pencils of the form $\mathrm{L}(\lambda) := A + \lambda B$, where $A, B \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathbb{C}$. Thus the set of n-by-n matrix pencils consists of affine transformations from $\mathbb{C}$ to $\mathbb{C}^{n\times n}$ which we denote by $\mathbb{P}_1(\mathbb{C}^{n\times n})$. Hence $\mathbb{P}_1(\mathbb{C}^{n\times n})$ is a vector space which we endow with an appropriate norm $|\!|\!|\cdot|\!|\!|$ as follows. Let $\mathrm{L} \in \mathbb{P}_1(\mathbb{C}^{n\times n})$ be given by $\mathrm{L}(\lambda) = A + \lambda B$. Then we define the pencil norm $|\!|\!|\mathrm{L}|\!|\!|$ by

$$|\!|\!|\mathrm{L}|\!|\!| := (\|A\|^2 + \|B\|^2)^{1/2}, \tag{3.1}$$

where $\|\cdot\|$ is either the spectral norm or the Frobenius norm on $\mathbb{C}^{n\times n}$. We refer to [2, 3] for various other norms on $\mathbb{P}_1(\mathbb{C}^{n\times n})$. It is evident that $\|\mathrm{L}(\lambda)\| \leq |\!|\!|\mathrm{L}|\!|\!|\ \|(1, \lambda)\|_2$.

The spectrum $\sigma(\mathrm{L})$ of a regular pencil $\mathrm{L} \in \mathbb{P}_1(\mathbb{C}^{n\times n})$ is given by

$$\sigma(\mathrm{L}) := \{\lambda \in \mathbb{C} : \mathrm{rank}(\mathrm{L}(\lambda)) < n\}.$$

To be precise, $\sigma(\mathrm{L})$ consists of finite eigenvalues of L. When B is singular, the pencil L has an infinite eigenvalue. In this chapter, we consider only finite eigenvalues of matrix pencils. By convention, if $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ then x is assumed to be nonzero, that is, $x \neq 0$. Treating (λ, x) as an approximate eigenpair of L, we define the backward error of (λ, x) by

$$\eta(\lambda, x, \mathrm{L}) := \inf\{|\!|\!|\triangle\mathrm{L}|\!|\!| : \triangle\mathrm{L} \in \mathbb{P}_1(\mathbb{C}^{n\times n}) \text{ and } \mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0\}.$$

We follow the convention that if L is given by $\mathrm{L}(\lambda) = A + \lambda B$ then the pencil $\triangle\mathrm{L}$ to be of the form $\triangle\mathrm{L}(\lambda) = \triangle A + \lambda\triangle B$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Then setting $r := -\mathrm{L}(\lambda)x$, we have

$$\eta(\lambda, x, \mathrm{L}) = \frac{\|r\|_2}{\|x\|_2\|(1, \lambda)\|_2}.$$

TH-789_04612301

Indeed, defining $\triangle A := \dfrac{rx^H}{x^H x(1+|\lambda|^2)}$ and $\triangle B := \dfrac{\overline{\lambda} r x^H}{x^H x(1+|\lambda|^2)}$, and considering the pencil $\triangle \mathrm{L}(z) = \triangle A + z\triangle B$, we have $|\!|\!|\triangle \mathrm{L}|\!|\!| = \|r\|_2/\|x\|_2\|(1,\lambda)\|_2$ and $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$.

Next, let $\mathbb{S}$ be a (real or complex) linear subspace of $\mathbb{P}_1(\mathbb{C}^{n\times n})$. Pencils in $\mathbb{S}$ will be referred to as structured pencils. Let $\mathrm{L} \in \mathbb{S}$. Then treating $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^{n\times n}$ as an approximate eigenpair of L, we define the structured backward error of (λ, x) by

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) := \inf\{|\!|\!|\triangle \mathrm{L}|\!|\!| : \triangle\mathrm{L} \in \mathbb{S} \text{ and } \mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0\}.$$

Obviously, we have $\eta(\lambda, x, \mathrm{L}) \leq \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$. Let L be given by $\mathrm{L}(z) = A + zB$. Then the ten special structures of L we consider in this chapter are as follows, where $* \in \{T, H\}$.

$\mathbb{S}$	Condition
T-symmetric	$A^T = A, B^T = B$
T-skew-symmetric	$A^T = -A, B^T = -B$
H-Hermitian	$A^H = A, B^H = B$
H-skew-Hermitian	$A^H = -A, B^H = -B$

$\mathbb{S}$	Condition
$*$-even	$A^* = A, B^* = -B$
$*$-odd	$A^* = -A, B^* = B$
$*$-palindromic	$A^* = B$

Let L be a regular pencil. We say that (λ, x, y) is an eigentriple of L if λ is an eigenvalue of L and x and y, respectively, are right and left eigenvectors of L corresponding to λ, that is, $\mathrm{L}(\lambda)x = 0$ and $y^H\mathrm{L}(\lambda) = 0$. An eigentriple (λ, x, y) is said to be normalized if $y^H y = x^H x = 1$. We consider only normalized eigentriples. Now, for ready reference, we collect some basic facts about eigenpairs of structured pencils in the following Theorem.

Theorem 3.2.1. *Let* $\mathrm{L} \in \mathbb{S}$ *be given by* $\mathrm{L}(z) = A + zB$. *Let* $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ *be an eigenpair of* L. *Then the results in Table 3.1 hold.*

$\mathbb{S}$	**eigenvalue pairing**	**eigentriple**	x^TAx	x^TBx
T-symmetric	λ	$(\lambda, x, \overline{x})$	in $\mathbb{C}$	in $\mathbb{C}$
T-skew-symmetric	λ	$(\lambda, x, \overline{x})$	0	0
T-even	$(\lambda, -\lambda)$	$(\lambda, x, \overline{y}), (-\lambda, y, \overline{x})$	0	0
T-odd	$(\lambda, -\lambda)$	$(\lambda, x, \overline{y}), (-\lambda, y, \overline{x})$	0	0 if $\lambda \neq 0$
T-palindromic	$(\lambda, 1/\lambda)$	$(\lambda, x, \overline{y}), (1/\lambda, y, \overline{x})$	0 if $\lambda \neq -1$	0 if $\lambda \neq -1$
	eigenvalue pairing	**eigentriple**	x^HAx	x^HBx
H-Hermitian / H-skew-Hermitian	$(\lambda, \overline{\lambda})$	(λ, x, y) $(\overline{\lambda}, y, x)$	0 if $\mathsf{im}(\lambda) \neq 0$	0 if $\mathsf{im}(\lambda) \neq 0$
H-even/ H-odd	$(\lambda, -\overline{\lambda})$	(λ, x, y) $(-\overline{\lambda}, y, x)$	0 if $\mathsf{re}(\lambda) \neq 0$	0 if $\mathsf{re}(\lambda) \neq 0$
H-palindromic	$(\lambda, 1/\overline{\lambda})$	$(\lambda, x, y), (1/\overline{\lambda}, y, x)$	0 if $|\lambda| \neq 1$	0 if $|\lambda| \neq 1$

Table 3.1: Eigensymmetry of structured pencils and related conditions.

Proof: Note that when L is T-symmetric or T-skew-symmetric, we have $\mathrm{L}(\lambda)x = 0$ and $\overline{x}^H\mathrm{L}(\lambda) = 0$. Hence $(\lambda, x, \overline{x})$ is an eigentriple of L. In particular, if L is T-skew-symmetric then both A and B are skew-symmetric and hence $x^TAx = 0 = x^TBx$.

TH-789_04612301

When L is T-even or T-odd, we have $\mathrm{L}(\lambda)^T = \mathrm{L}(-\lambda)$ or $\mathrm{L}(\lambda)^T = -\mathrm{L}(-\lambda)$. Hence if $\mathrm{L}(\lambda)x = 0$ and $\mathrm{L}(-\lambda)y = 0$ then $\overline{x}^H\mathrm{L}(-\lambda) = 0$ and $\overline{y}^H\mathrm{L}(\lambda) = 0$. This shows $(\lambda, -\lambda)$ pairing of eigenvalues and that $(\lambda, x, \overline{y})$ and $(-\lambda, y, \overline{x})$ are eigentriples. When L is T-even, B is skew-symmetric and hence $x^TBx = 0$. Consequently, $x^T\mathrm{L}(\lambda)x = 0 \Rightarrow x^TAx = 0$. Similarly, when L is T-odd, A is skew-symmetric and hence $x^TAx = 0$. Consequently, $x^T\mathrm{L}(\lambda)x = 0 \Rightarrow x^TBx = 0$ whenever $\lambda \neq 0$. The proof is similar for H-Hermitian, H-skew-Hermitian, H-odd and H-even pencils.

Now let L be T-palindromic given by $\mathrm{L}(z) = A + zA^T$. Suppose that $\lambda \neq 0$. Then $\mathrm{L}(\lambda)x = 0 \Rightarrow \overline{x}^H\mathrm{L}(1/\lambda) = 0$ which shows $(\lambda, 1/\lambda)$ pairing of eigenvalues. It also follows that $(\lambda, \overline{y}, x)$ is an eigentriple of L if and only if $(1/\lambda, \overline{x}, y)$ is an eigentriple of L. Note that $x^T\mathrm{L}(\lambda)x = 0 \Rightarrow x^TAx + \lambda x^TAx = 0$. Thus, if $\lambda \neq -1$ then $x^TAx = 0$.

Similarly, when L is H-palindromic it follows that (λ, y, x) is an eigentriple of L if and only if $(1/\overline{\lambda}, x, y)$ is an eigentriple of L. Now $x^H\mathrm{L}(\lambda)x \Rightarrow x^HAx + \lambda\,\overline{x^HAx} = 0 \Rightarrow |x^HAx| = |\lambda|\,|x^HAx| \Rightarrow (1-|\lambda|)\,|x^HAx| = 0$. Hence for $|\lambda| \neq 1$, we have $x^HAx = 0$. ■

Next, we show that if $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$ and $\mathrm{L} \in \mathbb{S}$ then there exists $\triangle\mathrm{L} \in \mathbb{S}$ such that (λ, x) is an eigenpair of $\mathrm{L} + \triangle\mathrm{L}$, that is, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$. Consequently, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) < \infty$.

Theorem 3.2.2. *Let $\mathbb{S} \in \{T$-symmetric, T-skew-symmetric, T-odd, T-even, H-Hermitian, H-skew-Hermitian, H-odd, H-even$\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zB$. Let $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$ be such that $x^Hx = 1$. Set $r := -\mathrm{L}(\lambda)x$ and define*

$$\triangle A := \begin{cases} -\overline{x}x^TAx x^H + \frac{1}{1+|\lambda|^2}[\overline{x}r^T + rx^H - 2(x^Tr)\overline{x}x^H], & \text{if } A = A^T, \\ -\frac{1}{1+|\lambda|^2}[\overline{x}r^T - rx^H], & \text{if } A = -A^T, \end{cases}$$

$$\triangle B := \begin{cases} -\overline{x}x^TBx x^H + \frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x}r^T + rx^H - 2(x^Tr)\overline{x}x^H], & \text{if } B = B^T, \\ -\frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x}r^T - rx^H], & \text{if } B = -B^T, \end{cases}$$

and

$$\triangle A := \begin{cases} -xx^HAxx^H + \frac{1}{1+|\lambda|^2}[xr^H(I - xx^H) + (I - xx^H)rx^H], & \text{if } A = A^H, \\ -xx^HAxx^H - \frac{1}{1+|\lambda|^2}[xr^H(I - xx^H) - (I - xx^H)rx^H], & \text{if } A = -A^H. \end{cases}$$

$$\triangle B := \begin{cases} -xx^HBxx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I - xx^H) + \overline{\lambda}(I - xx^H)rx^H] & \text{if } B = B^H \\ -xx^HBxx^H - \frac{1}{1+|\lambda|^2}[\lambda xr^H(I - xx^H) - \overline{\lambda}(I - xx^H)rx^H] & \text{if } B = -B^H. \end{cases}$$

Consider the pencil $\triangle\mathrm{L}(z) = \triangle A + z\triangle B$. Then $\triangle\mathrm{L} \in \mathbb{S}$ and $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$.

Proof: The proof is computational and is easy to check. ■.

For palindromic pencils, we have the following result.

Theorem 3.2.3. *Let $\mathbb{S} \in \{T$-palindromic, H-palindromic$\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zA^*$, where $A^* = A^T$ or $A^* = A^H$. Let $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$ be such that $x^Hx = 1$. Set $r := -\mathrm{L}(\lambda)x$ and define*

$$\triangle A := \begin{cases} -\overline{x}x^TAxx^H + \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H], & \text{if } B = A^T, \\ -xx^HAxx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I - xx^H) + (I - xx^H)rx^H], & \text{if } B = A^H. \end{cases}$$

TH-789_04612301

Consider the pencil $\triangle \mathrm{L}(z) = \triangle A + z(\triangle A)^*$. *Then* $\triangle \mathrm{L} \in \mathbb{S}$ *and* $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$.

Proof: The proof is computational and is easy to check. ■.

We also consider general classes of linearly structured pencils whose coefficient matrices are elements of certain Jordan and/or Lie algebras and show that for these pencils structured backward perturbation analysis ultimately reduces to that of one of the ten classes of structured pencils discussed above.

3.3 Frobenius norm and structured backward errors

Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Unless stated otherwise, we always assume that $x^H x = 1$. Let $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zB$. In this section, we determine the structured backward error $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ when $\mathbb{C}^{n \times n}$ is equipped with the Frobenius norm. Recall that the pencil norm defined in (3.1) is then given by $|||\mathrm{L}||| := \sqrt{\|A\|_F^2 + \|B\|_F^2} = \|[A \ B]\|_F$. Also recall that the unstructured backward error $\eta(\lambda, x, \mathrm{L})$ for the spectral norm as well as for the Frobenius norm on $\mathbb{C}^{n \times n}$ is given by $\eta(\lambda, x, \mathrm{L}) = \|\mathrm{L}(\lambda)x\|_2 / \|(1, \lambda)\|_2$.

Theorem 3.3.1. *Let* $\mathbb{S}$ *be the space of* T*-symmetric pencils and let* $\mathrm{L} \in \mathbb{S}$ *be given by* $\mathrm{L}(z) = A + zB$. *Then for* $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, *setting* $r := -\mathrm{L}(\lambda)x$, *we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \frac{\sqrt{2\|r\|_2^2 - |x^T r|^2}}{\|(1, \lambda)\|_2} \leq \sqrt{2}\, \eta(\lambda, x, \mathrm{L}).$$

Define $\triangle A := \dfrac{1}{1 + |\lambda|^2}[\overline{x} r^T + r x^H - (r^T x)\overline{x} x^H]$ *and* $\triangle B := \dfrac{\overline{\lambda}}{1 + |\lambda|^2}[\overline{x} r^T + r x^H - (r^T x)\overline{x} x^H]$ *and consider the pencil* $\triangle \mathrm{L}(z) = \triangle A + z \triangle B$. *Then* $\triangle \mathrm{L}$ *is* T*-symmetric,* $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$ *and* $|||\triangle \mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: By Theorem 3.2.2 there is a $\triangle \mathrm{L} \in \mathbb{S}$ such that $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$. Let $\triangle \mathrm{L}$ be given by $\triangle \mathrm{L}(z) = \triangle A + z \triangle B$. Then we have $(\triangle A + \lambda \triangle B)x = r$. Choose $Q_1 \in \mathbb{C}^{n \times (n-1)}$ such that $Q := [x, \ Q_1]$ is unitary. Then

$$\widetilde{\triangle A} := Q^T \triangle A Q = \begin{pmatrix} a_{11} & a_1^T \\ a_1 & A_1 \end{pmatrix}, \quad \widetilde{\triangle B} := Q^T \triangle B Q = \begin{pmatrix} b_{11} & b_1^T \\ b_1 & B_1 \end{pmatrix}, \quad Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix},$$

where $A_1 = A_1^T$ and $B_1 = B_1^T$ are of size $n-1$. Since $\overline{Q} Q^T = I$, we have

$$(\overline{Q} \widetilde{\triangle A} Q^H + \lambda \overline{Q} \widetilde{\triangle B} Q^H)x = r \Rightarrow (\widetilde{\triangle A} Q^H + \lambda \widetilde{\triangle B} Q^H)x = Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

As $Q^H x = e_1$, the first column of the identity matrix, we have

$$(\widetilde{\triangle A} + \lambda \widetilde{\triangle B}) Q^H x = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} + \lambda b_{11} \\ a_1 + \lambda b_1 \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

TH-789_04612301

This gives $a_{11} + \lambda b_{11} = x^T r$ and $a_1 + \lambda b_1 = Q_1^T r$ whose minimum norm solutions are

$$(a_1 \;\; b_1) = Q_1^T r \begin{pmatrix} 1 \\ \lambda \end{pmatrix}^{\dagger} \Rightarrow a_1 = \frac{1}{1+|\lambda|^2} Q_1^T r, \;\; b_1 = \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r$$

and $(a_{11} \;\; b_{11}) = x^T r \begin{pmatrix} 1 \\ \lambda \end{pmatrix}^{\dagger} \Rightarrow a_{11} = \frac{1}{1+|\lambda|^2} x^T r, \;\; b_{11} = \frac{\overline{\lambda}}{1+|\lambda|^2} x^T r$. Hence we have

$$\widetilde{\triangle A} = \begin{pmatrix} \frac{1}{1+|\lambda|^2} x^T r & \frac{1}{1+|\lambda|^2} (Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix}, \;\; \widetilde{\triangle B} = \begin{pmatrix} \frac{\overline{\lambda}}{1+|\lambda|^2} x^T r & \frac{\overline{\lambda}}{1+|\lambda|^2} (Q_1^T r)^T \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r & B_1 \end{pmatrix}.$$

This shows that the Frobenius norms of $\widetilde{\triangle A}$ and $\widetilde{\triangle B}$ are minimized when $A_1 = 0$ and $B_1 = 0$. Hence $\|\triangle A\|_F^2 = \|\widetilde{\triangle A}\|_F^2 = |a_{11}|^2 + 2\,\|a_1\|_2^2$ and $\|\triangle B\|_F^2 = \|\widetilde{\triangle B}\|_F^2 = |b_{11}|^2 + 2\,\|b_1\|_2^2$. Note that $QQ^H = I \Rightarrow Q_1 Q_1^H = I - xx^H \Rightarrow \overline{Q}_1 Q_1^T = I - \overline{x}x^T$. Consequently, we have

$$\|\triangle \mathrm{L}\| = (\|\triangle A\|_F^2 + \|\triangle B\|_F^2)^{1/2} = \frac{\sqrt{|x^T r|^2 + 2\,\|(I - \overline{x}x^T)r\|_2^2}}{\|(1,\lambda)\|_2} = \frac{\sqrt{2\|r\|_2^2 - |x^T r|^2}}{\|(1,\lambda)\|_2}.$$

Next, we have

$$\begin{aligned} \triangle A &= \overline{Q}\widetilde{\triangle A} Q^H = \frac{1}{1+|\lambda|^2} \overline{x} x^T r x^H + \frac{1}{1+|\lambda|^2} [\overline{x} r^T Q_1 Q_1^H + \overline{Q}_1 Q_1^T r x^H] + \overline{Q}_1 A_1 Q_1^H \\ &= \frac{1}{1+|\lambda|^2} [\overline{x} r^T + r x^H - (r^T x) \overline{x} x^H] + \overline{Q}_1 A_1 Q_1^H, \\ \triangle B &= \overline{Q}\widetilde{\triangle B} Q^H = \frac{\overline{\lambda}}{1+|\lambda|^2} \overline{x} x^T r x^H + \frac{1}{1+|\lambda|^2} [\overline{\lambda} \overline{x} r^T Q_1 Q_1^H + \overline{\lambda Q}_1 Q_1^T r x^H] + \overline{Q}_1 B_1 Q_1^H \\ &= \frac{\overline{\lambda}}{1+|\lambda|^2} [\overline{x} r^T + r x^H - (r^T x) \overline{x} x^H] + \overline{Q}_1 B_1 Q_1^H \end{aligned}$$

from which we obtain the desired pencil by setting $A_1 = 0$ and $B_1 = 0$. This completes the proof. ■

Observe that if Y is symmetric and $Yx = 0$ then $Y = (I - xx^H)^T Z (I - xx^H)$ for some symmetric matrix Z. Consequently, we have $\overline{Q_1} A_1 Q_1^H = (I - xx^H)^T Z_1 (I - xx^H)$ and $\overline{Q_1} B_1 Q_1^H = (I - xx^H)^T Z_2 (I - xx^H)$ for some symmetric matrices Z_1 and Z_2. Hence from the proof of Theorem 3.3.1 we have following.

Corollary 3.3.2. *Let* L *be a* T*-symmetric pencil and* $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. *Set* $r := -\mathrm{L}(\lambda)x$. *Let* K *be a* T*-symmetric pencil. Then* $\mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0$ *if and only if* $\mathrm{K}(z) = \triangle\mathrm{L}(z) + (I - xx^H)^T \mathrm{N}(z)(I - xx^H)$ *for some* T*-symmetric pencil* N, *where* $\triangle\mathrm{L}$ *is the* T*-symmetric pencil given in Theorem 3.3.1.*

Next, we consider T-skew-symmetric pencils.

Theorem 3.3.3. *Let* $\mathbb{S}$ *be the space of* T*-skew-symmetric pencils and let* $\mathrm{L} \in \mathbb{S}$ *be given by* $\mathrm{L}(z) = A + zB$. *Let* $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ *and* $r := -\mathrm{L}(\lambda)x$. *Then* $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{2}\,\|r\|_2 / \|(1,\lambda)\|_2 = \sqrt{2}\,\eta(\lambda, x, \mathrm{L})$. *Further, for the* T*-skew-symmetric pencil* $\triangle\mathrm{L}$ *given in Theorem 3.2.2, we have* $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

TH-789_04612301

Proof: As A and B are skew-symmetric, from the proof of Theorem 3.3.1, we have

$$\widetilde{\triangle A} = Q^T \triangle A Q = \begin{pmatrix} 0 & a_1^T \\ -a_1 & A_1 \end{pmatrix}, \quad \widetilde{\triangle B} = Q^T \triangle B Q = \begin{pmatrix} 0 & b_1^T \\ -b_1 & B_1 \end{pmatrix}, \quad Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix},$$

where A_1 and B_1 are skew-symmetric matrices of size $n-1$. Consequently, as before, we have $(\widetilde{\triangle A} + \lambda \widetilde{\triangle B}) Q^H x = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}$ which gives $\begin{pmatrix} 0 \\ -a_1 - \lambda b_1 \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}$. Note that $x^T r = 0$ and the smallest norm solution of $-a_1 - \lambda b_1 = Q_1^T r$ is given by

$$(a_1 \;\; b_1) = Q_1^T r \begin{pmatrix} -1 \\ -\lambda \end{pmatrix}^{\dagger} \Rightarrow a_1 = -\frac{1}{1+|\lambda|^2} Q_1^T r, \;\; b_1 = -\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r.$$

Hence we have

$$\triangle A = \overline{Q} \begin{pmatrix} 0 & -\frac{1}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix} Q^H, \;\; \triangle B = \overline{Q} \begin{pmatrix} 0 & -\frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r & B_1 \end{pmatrix} Q^H.$$

Setting $A_1 = 0$ and $B_1 = 0$ we obtain $\triangle \mathrm{L}$ such that $\|\triangle \mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{2}\|r\|_2 / \|(1, \lambda)\|_2$.

Since $\overline{Q}_1 Q_1^T = I - \overline{x} x^T$, we have

$$\triangle A = -\frac{1}{1+|\lambda|^2}[\overline{x} r^T - r x^H] + \overline{Q}_1 A_1 Q_1^H \text{ and } \triangle B = -\frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x} r^T - r x^H] + \overline{Q}_1 B_1 Q_1^H.$$

Setting $A_1 = B_1 = 0$ we obtain the T-skew-symmetric pencil $\triangle \mathrm{L}$ given in Theorem 3.2.2. ■

Using the fact that if Y is skew-symmetric and $Yx = 0$ then $Y = (I - xx^H)^T Z (I - xx^H)$ for some skew-symmetric matrix Z, we obtain an analogue of Corollary 3.3.2 for T-skew-symmetric pencils.

Next, we derive structured backward errors for T-even and T-odd pencils.

Theorem 3.3.4. *Let $\mathbb{S} \in \{T\text{-even}, T\text{-odd}\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zB$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{|x^T A x|^2 + \frac{2\|r\|_2^2 - 2|x^T r|^2}{1+|\lambda|^2}} = \frac{\sqrt{2\|r\|_2^2 + (|\lambda|^2 - 1)|x^T r|^2}}{\|(1, \lambda)\|_2}$$

when L is T-even and

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{|x^T B x|^2 + \frac{2\|r\|_2^2 - 2|x^T r|^2}{1+|\lambda|^2}} = \frac{\sqrt{2\|r\|_2^2 + (|\lambda|^{-2} - 1)|x^T r|^2}}{\|(1, \lambda)\|_2}, & \text{if } \lambda \neq 0. \\ \sqrt{2}\,\eta(\lambda, x, \mathrm{L}), & \text{if } \lambda = 0, \end{cases}$$

when L is T-odd. The pencil $\triangle \mathrm{L} \in \mathbb{S}$ given in Theorem 3.2.2 satisfies $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$ and $\|\triangle \mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: First, assume that L is T-even. Then noting that $A = A^T$ and $B = -B^T$, the proof follows from similar arguments as those employed for T-symmetric and T-skew-symmetric

TH-789_04612301

pencils. Indeed, considering a unitary matrix $Q := [x,\ Q_1]$, we have

$$\widetilde{\triangle A} := Q^T \triangle A Q = \begin{pmatrix} a_{11} & a_1^T \\ a_1 & A_1 \end{pmatrix}, \quad \widetilde{\triangle B} := Q^T \triangle B Q = \begin{pmatrix} 0 & b_1^T \\ -b_1 & B_1 \end{pmatrix}, \quad Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix},$$

where $A_1 = A_1^T$ and $B_1 = -B_1^T$ are of size $n-1$. Consequently, we have $(\widetilde{\triangle A} + \lambda \widetilde{\triangle B}) Q^H x = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} \\ a_1 - \lambda b_1 \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}$. This gives $a_{11} = -x^T A x$. The smallest norm solution of $a_1 + \lambda b_1 = Q_1^T r$ is given by

$$(a_1\ \ b_1) = Q_1^T r \begin{pmatrix} 1 \\ -\lambda \end{pmatrix}^{\dagger} \Rightarrow a_1 = \frac{1}{1+|\lambda|^2} Q_1^T r, \quad b_1 = -\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r.$$

Consequently, we have

$$\triangle A = \overline{Q} \begin{pmatrix} -x^T A x & (\frac{1}{1+|\lambda|^2} Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix} Q^H, \ \triangle B = \overline{Q} \begin{pmatrix} 0 & (-\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r)^T \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r & B_1 \end{pmatrix} Q^H.$$

Setting $A_1 = B_1 = 0$ and using the fact that $\overline{Q}_1 Q_1^T = I - \overline{x} x^T$, we obtain the pencil $\triangle \mathrm{L}$ such that

$$|||\triangle \mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{|x^T A x|^2 + \frac{2\|r\|_2^2 - 2|x^T r|^2}{1+|\lambda|^2}}.$$

Now simplifying expressions for $\triangle A$ and $\triangle B$, we obtain

$$\begin{aligned} \triangle A &= -\overline{x} x^T A x x^H + \frac{1}{1+|\lambda|^2}[\overline{x} r^T + r x^H - 2(x^T r)\overline{x} x^H] + \overline{Q}_1 A_1 Q_1^H, \\ \triangle B &= -\frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x} r^T - r x^H] + \overline{Q}_1 B_1 Q_1^H. \end{aligned}$$

Setting $A_1 = B_1 = 0$ we obtain the T-even pencil $\triangle \mathrm{L}$ given in Theorem 3.2.2.

When L is T-odd, the results follow by interchanging the role of A and B. ■

It follows from Theorem 3.3.4 that for a T-even pencil, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \sqrt{2}\, \eta(\lambda, x, \mathrm{L})$ when $|\lambda| \leq 1$ and $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \|(1, \lambda)\|_2\, \eta(\lambda, x, \mathrm{L})$ when $|\lambda| > 1$. Similarly, for a T-odd pencil, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \sqrt{2}\, \eta(\lambda, x, \mathrm{L})$ when $|\lambda| \geq 1$ and $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \|(1, \lambda^{-1})\|_2\, \eta(\lambda, x, \mathrm{L})$ when $\lambda \neq 0$ and $|\lambda| < 1$.

We mention that an analogue of Corollary 3.3.2 holds for T-even and T-odd pencils as well. Now, we consider a T-palindromic pencil $\mathrm{L}(z) = A + zA^T$.

Theorem 3.3.5. *Let $\mathbb{S}$ be the space of T-palindromic pencils and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zA^T$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\sqrt{|x^T A x|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{1+|\lambda|^2}} = \sqrt{2}\, \frac{\sqrt{\|r\|_2^2 - 2\mathrm{re}\lambda\, |x^T A x|^2}}{\|(1,\ \lambda)\|_2}, & \text{if } \lambda \neq -1, \\ \sqrt{2}\ \eta(\lambda, x, \mathrm{L}), & \text{if } \lambda = -1, \end{cases}$$

In particular, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{2}\ \eta(\lambda, x, \mathrm{L})$, if $\lambda \in i\mathbb{R}$.

TH-789_04612301

Now define

$$\triangle A = \begin{cases} \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H], & \text{if } \lambda = -1, \\ -\overline{x}x^T Axx^H + \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H], & \text{if } \lambda \neq -1, \end{cases}$$

and consider the pencil $\triangle \mathrm{L}(z) = \triangle A + z(\triangle A)^T$. *Then* $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$ *and* $|||\triangle \mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: By Theorem 3.2.3, there exists a T-palindromic pencil $\triangle \mathrm{L}(z) = \triangle A + z\triangle A^T$ such that $(\mathrm{L}(\lambda) + \triangle \mathrm{L}(\lambda))x = 0$. Let $Q_1 \in \mathbb{C}^{n\times(n-1)}$ be such that $Q := [x \;\; Q_1]$ is unitary. Then

$$\widetilde{\triangle A} := Q^T \triangle A Q = \begin{pmatrix} a_{11} & a_1^T \\ b_1 & A_1 \end{pmatrix}, \quad Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

Now, if $\lambda \neq -1$ then by Theorem 3.2.1, we have $x^T(\triangle A + A)x = 0 \Rightarrow x^T \triangle A x = -x^T Ax$. Hence we have $a_{11} = -x^T Ax$. When $\lambda = -1$, we have $\lambda a_{11}^T + a_{11} = x^T r = 0$ for any a_{11}. Since the aim is to minimize the Frobenius norm of $\triangle A$, we set $a_{11} = 0$.

Next, the minimum norm solution of $a_1\lambda + b_1 = Q_1^T r$ is given by

$$\begin{pmatrix} a_1 & b_1 \end{pmatrix} = Q_1^T r \begin{pmatrix} \lambda \\ 1 \end{pmatrix}^{\dagger} \Rightarrow a_1 = \frac{\overline{\lambda} Q_1^T r}{1 + |\lambda|^2}, \quad b_1 = \frac{Q_1^T r}{1 + |\lambda|^2}.$$

Therefore when $\lambda = -1$, we have $\triangle A = \overline{Q} \begin{pmatrix} 0 & (\frac{\overline{\lambda} Q_1^T r}{1+|\lambda|^2})^T \\ \frac{Q_1^T r}{1+|\lambda|^2} & A_1 \end{pmatrix} Q^H$. Setting $A_1 = 0$, we obtain $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = |||\triangle \mathrm{L}||| = \sqrt{2}\|r\|_2/\sqrt{1 + |\lambda|^2} = \sqrt{2}\,\eta(\lambda, x, \mathrm{L})$. Since $Q_1 Q_1^H = I - xx^H \Rightarrow \overline{Q}_1 Q_1^T = I - \overline{x}x^T$, simplifying the expression for $\triangle A$, we obtain

$$\triangle A = \frac{1}{1 + |\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H] + \overline{Q}_1 A_1 Q_1^H.$$

When $\lambda \neq -1$, we have $\triangle A = \overline{Q} \begin{pmatrix} -x^T Ax & (\frac{\overline{\lambda} Q_1^T r}{1+|\lambda|^2})^T \\ \frac{Q_1^T r}{1+|\lambda|^2} & A_1 \end{pmatrix} Q^H$. Setting $A_1 = 0$, we obtain

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = |||\triangle \mathrm{L}||| = \sqrt{2|x^T Ax|^2 + \frac{2\|[I - \overline{x}x^T]r\|_2^2}{1 + |\lambda|^2}} = \sqrt{2}\sqrt{|x^T Ax|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{1 + |\lambda|^2}}$$

from which the result follows. Since $|x^T r|^2 = |x^T Ax|^2(1 + |\lambda|^2)$ when $\lambda \in i\mathbb{R}$, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{2}\|r\|_2/\|(1, \lambda)\|_2$, for $\lambda \in i\mathbb{R}$. Again, simplifying the expression for $\triangle A$, we obtain

$$\triangle A = -\overline{x}x^T Axx^H + \frac{1}{1 + |\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H] + \overline{Q}_1 A_1 Q_1^H.$$

This completes the proof. ■

Observe that by Theorem 3.3.5 we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \sqrt{2}\,\eta(\lambda, x, \mathrm{L})$ when $\mathsf{re}\lambda > 0$ and $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \|(1, \sqrt{|\mathsf{re}\lambda|}/|1 + \lambda|)\|_2\,\eta(\lambda, x, \mathrm{L})$ when $\lambda \neq -1$ and $\mathsf{re}\lambda < 0$.

TH-789_04612301

Note that if $Y \in \mathbb{C}^{n\times n}$ is such that $Yx = 0$ and $Y^T x = 0$ then $Y = (I - xx^H)^T Z(I - xx^H)$ for some matrix Z. Hence from the proof of Theorem 3.3.5, we obtain an analogue of Corollary 3.3.2 for T-palindromic pencil. Indeed, if K is a T-palindromic pencil such that $\mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K}(z) = \triangle\mathrm{L}(z) + (I - xx^H)^T \mathrm{N}(z)(I - xx^H)$ for some T-palindromic pencil N, where $\triangle\mathrm{L}$ is given in Theorem 3.3.5.

Now we turn to H-Hermitian, H-skew-Hermitian, H-even, H-odd and H-palindromic pencils.

Theorem 3.3.6. *Let $\mathbb{S} \in \{H\text{-Hermitian}, H\text{-skew-Hermitian}\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zB$. For $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, set $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \dfrac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|(1,\lambda)\|_2} \leq \sqrt{2}\,\eta(\lambda, x, \mathrm{L}) & \text{if } \lambda \in \mathbb{R}, \\ \sqrt{|x^H A x|^2 + |x^H B x|^2 + \dfrac{2\|r\|_2^2 - 2|x^H r|^2}{1 + |\lambda|^2}}, & \text{if } \lambda \in \mathbb{C}\setminus\mathbb{R}. \end{cases}$$

In particular, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \|r\|_2 = \sqrt{2}\,\eta(\lambda, x, \mathrm{L})$, if $\lambda = \pm i$.

When $\lambda \in \mathbb{R}$, define

$$\triangle A := \begin{cases} \frac{1}{1+\lambda^2}[xr^H + rx^H - (r^H x)xx^H], & \text{if } A = A^H \\ \frac{1}{1+\lambda^2}[rx^H - xr^H + (r^H x)xx^H], & \text{if } A = -A^H \end{cases}$$

$$\triangle B := \begin{cases} \frac{\lambda}{1+\lambda^2}[xr^H + rx^H - (r^H x)xx^H], & \text{if } B = B^H \\ \frac{\lambda}{1+\lambda^2}[rx^H - xr^H + (r^H x)xx^H], & \text{if } B = -B^H \end{cases}$$

and consider the pencil $\triangle\mathrm{L}(z) = \triangle A + z\triangle B$. Then $\triangle\mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ and $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

When $\lambda \in \mathbb{C}\setminus\mathbb{R}$, the H-Hermitian/H-skew-Hermitian pencil $\triangle\mathrm{L}$ given in Theorem 3.2.2 satisfies $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ and $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: Suppose that $\mathrm{L}(z) = A + zB$ is H-Hermitian so that $A = A^H$ and $B = B^H$. By Theorem 3.2.2 there exists H-Hermitian pencil $\triangle\mathrm{L}(z) = \triangle A + z\triangle B$ such that $(\triangle A + \lambda\triangle B)x = r$. Again, choosing a unitary matrix $Q := [x,\ Q_1]$, we have

$$\widetilde{\triangle A} := Q^H \triangle A Q = \begin{pmatrix} a_{11} & a_1^H \\ a_1 & A_1 \end{pmatrix}, \quad \widetilde{\triangle B} := Q^H \triangle B Q = \begin{pmatrix} b_{11} & b_1^H \\ b_1 & B_1 \end{pmatrix}, \quad Q^H r = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix},$$

where $A_1 = A_1^H$ and $B_1 = B_1^H$ are of size $n-1$. This gives $(\widetilde{\triangle A} + \lambda\widetilde{\triangle B})Q^H x = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix} \Rightarrow$ $\begin{pmatrix} a_{11} + \lambda b_{11} \\ a_1 + \lambda b_1 \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. The minimum norm solution of $a_1 + \lambda b_1 = Q_1^H r$ is given by

$$(a_1 \;\; b_1) = Q_1^H r \begin{pmatrix} 1 \\ \lambda \end{pmatrix}^{\dagger} \Rightarrow a_1 = \frac{1}{1+|\lambda|^2} Q_1^H r, \quad b_1 = \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r.$$

For the equation $a_{11} + \lambda b_{11} = x^H r$, two cases arise.

TH-789_04612301

Case-I: When $\lambda \in \mathbb{R}$, the minimum norm solution is given by

$$(a_{11} \;\; b_{11}) = x^H r \begin{pmatrix} 1 \\ \lambda \end{pmatrix}^{\dagger} \Rightarrow a_{11} = \frac{1}{1+\lambda^2} x^H r \in \mathbb{R}, \;\; b_{11} = \frac{\lambda}{1+\lambda^2} x^H r \in \mathbb{R}.$$

Hence we have

$$\triangle A = Q \begin{pmatrix} \frac{1}{1+\lambda^2} x^H r & \frac{1}{1+\lambda^2} (Q_1^H r)^H \\ \frac{1}{1+\lambda^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \; \triangle B = Q \begin{pmatrix} \frac{\lambda}{1+\lambda^2} x^H r & (\frac{\lambda}{1+\lambda^2} Q_1^H r)^H \\ \frac{\lambda}{1+\lambda^2} Q_1^H r & B_1 \end{pmatrix} Q^H.$$

Setting $A_1 = B_1 = 0$ and using the fact that $Q_1 Q_1^H = I - xx^H$, we have

$$\eta^{\mathbb{S}}(x, \lambda, \mathrm{L}) = \|\triangle \mathrm{L}\| = \frac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|(1,\lambda)\|_2}.$$

Now simplifying the expressions for $\triangle A$ and $\triangle B$, we have

$$\begin{aligned} \triangle A &= \frac{1}{1+\lambda^2} x x^H r x^H + \frac{1}{1+\lambda^2}[x r^H Q_1 Q_1^H + Q_1 Q_1^H r x^H] + Q_1 A_1 Q_1^H \\ &= \frac{1}{1+\lambda^2}[x r^H + r x^H - (r^H x) x x^H] + Q_1 A_1 Q_1^H, \\ \triangle B &= \frac{\lambda}{1+\lambda^2} x x^H r x^H + \frac{\lambda}{1+\lambda^2}[x r^H Q_1 Q_1^H + Q_1 Q_1^H r x^H] + Q_1 B_1 Q_1^H \\ &= \frac{\lambda}{1+\lambda^2}[x r^H + r x^H - (r^H x) x x^H] + Q_1 B_1 Q_1^H. \end{aligned}$$

Hence the results follow.

Case-II: Suppose that $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then by Theorem 3.2.1, we have $x^H(A + \triangle A)x = 0$ and $x^H(B + \triangle B)x = 0$. Hence we have $a_{11} = -x^H A x$ and $b_{11} = -x^H B x$. Consequently,

$$\triangle A = Q \begin{pmatrix} -x^H A x & (\frac{1}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \;\; \triangle B = Q \begin{pmatrix} -x^H B x & (\frac{\lambda}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{\lambda}{1+|\lambda|^2} Q_1^H r & B_1 \end{pmatrix} Q^H.$$

Setting $A_1 = B_1 = 0$, we obtain

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \|\triangle \mathrm{L}\| = \sqrt{|x^H A x|^2 + |x^H B x|^2 + \frac{2\|(I - xx^H) r\|_2^2}{1+|\lambda|^2}}.$$

Hence the result follows.

Now simplifying the expressions for $\triangle A$ and $\triangle B$, we have

$$\begin{aligned} \triangle A &= -x x^H A x x^H + \frac{1}{1+|\lambda|^2}[x r^H (I - x x^H) + (I - x x^H) r x^H] + Q_1 A_1 Q_1^H, \\ \triangle B &= -x x^H B x x^H + \frac{1}{1+|\lambda|^2}[\lambda x r^H (I - x x^H) + \overline{\lambda}(I - x x^H) r x^H] + Q_1 B_1 Q_1^H. \end{aligned}$$

Setting $A_1 = B_1 = 0$, we obtain the H-Hermitian pencil $\triangle \mathrm{L}$ given in Theorem 3.2.2.

The proof is similar for the case when L is H-skew-Hermitian. ■

Needless to mention that an analogue of Corollary 3.3.2 holds for H-Hermitian/H-skew-Hermitian pencils.

TH-789_04612301

Theorem 3.3.7. *Let* $\mathbb{S} \in \{H\text{-even}, H\text{-odd}\}$ *and* $\mathrm{L} \in \mathbb{S}$ *be given by* $\mathrm{L}(z) = A + zB$. *For* $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, *set* $r := -\mathrm{L}(\lambda)x$. *Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \dfrac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|(1,\, \lambda)\|_2} \leq \sqrt{2}\,\eta(\lambda, x, \mathrm{L}) & \text{if } \lambda \in i\mathbb{R}, \\ \sqrt{|x^H A x|^2 + |x^H B x|^2 + \dfrac{2\|r\|_2^2 - 2|x^H r|^2}{1 + |\lambda|^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\mathbb{R}. \end{cases}$$

In particular, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \|r\|_2 = \sqrt{2}\ \eta(\lambda, x, \mathrm{L})$, *if* $\lambda = \pm 1$.

When $\lambda \in i\mathbb{R}$, *define*

$$\triangle A := \begin{cases} \frac{1}{1+|\lambda|^2}[xr^H + rx^H - (r^H x)xx^H], & \text{if } A = A^H \\ \frac{1}{1+|\lambda|^2}[rx^H - xr^H + (r^H x)xx^H], & \text{if } A = -A^H \end{cases}$$

$$\triangle B := \begin{cases} \frac{-\lambda}{1+|\lambda|^2}[rx^H - xr^H + (r^H x)xx^H], & \text{if } B = B^H \\ \frac{-\lambda}{1+|\lambda|^2}[rx^H + xr^H - (r^H x)xx^H], & \text{if } B = -B^H \end{cases}$$

and consider the pencil $\triangle\mathrm{L}(z) = \triangle A + z\triangle B$. *Then* $\triangle\mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

When $\lambda \in \mathbb{C} \setminus i\mathbb{R}$, *the* H*-even/*H*-odd pencil* $\triangle\mathrm{L}$ *given in Theorem 3.2.2 satisfies* $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: First, suppose that $\mathrm{L}(z) = A + zB$ is H even. Then $A = A^H$ and $B = -B^H$. By Theorem 3.2.2 there exists H-even pencil $\triangle\mathrm{L}(z) = \triangle A + z\triangle B$ such that $\triangle\mathrm{L}(\lambda)x = r$. Now choosing a unitary matrix $Q := [x, Q_1]$ and noting that $\triangle A = \triangle A^H$, $\triangle B = -\triangle B^H$, we have

$$\triangle A := Q \begin{pmatrix} a_{11} & a_1^H \\ a_1 & A_1 \end{pmatrix} Q^H \text{ and } \triangle B = Q \begin{pmatrix} b_{11} & b_1^H \\ -b_1 & B_1 \end{pmatrix} Q^H,$$

where $A_1 = A_1^H$ and $B_1 = -B_1^H$ are matrices of size $n-1$. Then $\triangle\mathrm{L}(\lambda)x = r$ gives $\begin{pmatrix} a_{11} + \lambda b_{11} \\ a_1 - \lambda b_1 \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. The minimum norm solution of $a_1 - \lambda b_1 = Q_1^H r$ is given by

$$(a_1 \ \ b_1) = Q_1^H r \begin{pmatrix} 1 \\ -\lambda \end{pmatrix}^{\dagger} \Rightarrow a_1 = \frac{1}{1+|\lambda|^2} Q_1^H r, \ \ b_1 = -\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r.$$

For the solution of $a_{11} + \lambda b_{11} = x^H r$ two cases arise. When $\lambda \in i\mathbb{R}$, the minimum norm solution is given by

$$(a_{11} \ \ b_{11}) = x^H r \begin{pmatrix} 1 \\ \lambda \end{pmatrix}^{\dagger} \Rightarrow a_{11} = \frac{1}{1+|\lambda|^2} x^H r \in \mathbb{R}, \ \ b_{11} = \frac{\overline{\lambda}}{1+|\lambda|^2} x^H r \in i\mathbb{R}.$$

When $\lambda \in \mathbb{C} \setminus i\mathbb{R}$, by Theorem 3.2.1, $x^H(A + \triangle A)x = 0 = x^H(B + \triangle B)x \Rightarrow a_{11} = -x^H A x$ and $b_{11} = -x^H B x$. Consequently, we have

$$\triangle A = Q \begin{pmatrix} \frac{1}{1+|\lambda|^2} x^H r & (\frac{1}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \triangle B = Q \begin{pmatrix} \frac{\overline{\lambda}}{1+|\lambda|^2} x^H r & (-\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r & B_1 \end{pmatrix} Q^H$$

TH-789_04612301

when $\lambda \in i\mathbb{R}$ and

$$\triangle A = Q \begin{pmatrix} -x^H A x & (\frac{1}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \ \triangle B = Q \begin{pmatrix} -x^H B x & (-\frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r)^H \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r & B_1 \end{pmatrix} Q^H$$

when $\lambda \in \mathbb{C} \setminus i\mathbb{R}$. Hence the desired results follow. Finally, reversing the role of A and B we obtain the results for the case when $\mathrm{L}(z) = A + zB$ is H-odd. ■

We have the following result for H-palindromic pencils.

Theorem 3.3.8. *Let $\mathbb{S}$ be the space of H-palindromic pencils and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) = A + zA^H$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\sqrt{|x^H A x|^2 + \dfrac{\|r\|_2^2 - |x^H r|^2}{1+|\lambda|^2}} & \text{if } |\lambda| \neq 1, \\ \sqrt{\|r\|_2^2 - \frac{1}{2}|x^H r|^2}, & \text{if } |\lambda| = 1. \end{cases}$$

Now define

$$\triangle A := \begin{cases} \frac{1}{1+|\lambda|^2}[r x^H + \lambda x r^H (I - x x^H)], & \text{if } |\lambda| = 1, \\ -x x^H A x x^H + \frac{1}{1+|\lambda|^2}[\lambda x r^H (I - x x^H) + (I - x x^H) r x^H], & \text{if } |\lambda| \neq 1, \end{cases}$$

and consider $\triangle \mathrm{L}(z) := \triangle A + z(\triangle A)^H$. Then $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ and $\|\triangle \mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: Let $Q := [x, \ \ Q_1]$ be unitary. Then $\widetilde{\triangle A} := Q^H \triangle A Q = \begin{pmatrix} a_{11} & a_1^H \\ b_1 & A_1 \end{pmatrix}$ and $Q^H r = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. Hence $\triangle\mathrm{L}(\lambda)x = r$ gives $\begin{pmatrix} \lambda a_{11}^H + a_{11} \\ \lambda a_1 + b_1 \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. If $|\lambda| \neq 1$ then by Theorem 3.2.1, we have $x^H(\triangle A + A)x = 0 \Rightarrow x^H \triangle A x = -x^H A x$. Hence we have $a_{11} = -x^H A x$. On the other hand, when $|\lambda| = 1$, the minimum norm solution is given by

$$(\overline{a}_{11} \ \ a_{11}) = x^H r \begin{pmatrix} \lambda \\ 1 \end{pmatrix}^{\dagger} = (\frac{\overline{\lambda} x^H r}{1+|\lambda|^2} \ \ \frac{x^H r}{1+|\lambda|^2}).$$

Note that when $|\lambda| = 1$ we have $\overline{x^H r} = \overline{\lambda} x^H r$. Next, the minimum solution of $a_1 \lambda + b_1 = Q_1^H r$ is given by $(a_1, \ \ b_1) = Q_1^H r \begin{pmatrix} \lambda \\ 1 \end{pmatrix}^{\dagger} = \begin{pmatrix} \frac{\overline{\lambda} Q_1^H r}{1+|\lambda|^2} & \frac{Q_1^H r}{1+|\lambda|^2} \end{pmatrix}$. Consequently, when $|\lambda| \neq 1$, we have $\triangle A = Q \begin{pmatrix} -x^H A x & (\frac{\overline{\lambda} Q_1^H r}{1+|\lambda|^2})^H \\ \frac{Q_1^H r}{1+|\lambda|^2} & A_1 \end{pmatrix} Q^H$. Setting $A_1 = 0$, we obtain

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \|\triangle \mathrm{L}\| = \sqrt{2|x^H A x|^2 + \frac{2\|[I - x x^H] r\|_2^2}{1+|\lambda|^2}} = \sqrt{2}\sqrt{|x^H A x|^2 + \frac{\|r\|_2^2 - |r^H x|^2}{1+|\lambda|^2}}.$$

Using the fact that $Q_1 Q_1^H = I - x x^H$, we have

$$\triangle A = -x x^H A x x^H + \frac{1}{1+|\lambda|^2}[\lambda x r^H (I - x x^H) + (I - x x^H) r x^H] + Q_1 A_1 Q_1^H.$$

TH-789_04612301

Setting $A_1 = 0$, the result follows.

For the case when $|\lambda| = 1$, we have $\triangle A = Q \begin{pmatrix} \frac{x^H r}{1+|\lambda|^2} & (\frac{\overline{\lambda} Q_1^H r}{1+|\lambda|^2})^H \\ \frac{Q_1^H r}{1+|\lambda|^2} & A_1 \end{pmatrix} Q^H$. Again setting $A_1 = 0$ we obtain

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = |||\triangle \mathrm{L}||| = \sqrt{\|r\|_2^2 - \frac{1}{2}|x^H r|^2}.$$

Since $Q_1 Q_1^H = (I - xx^H)$, simplifying the expression for $\triangle A$, we obtain

$$\triangle A := \frac{1}{1+|\lambda|^2}[rx^H + \lambda x r^H (I - xx^H)] + Q_1 A_1 Q_1^H.$$

Hence the proof. ■

Remark 3.3.9. *Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ with $x^H x = 1$ and $\mathbb{S} \in \{T$-symmetric, T-skew-symmetric, T-odd, T-even, T-palindromic, H-Hermitian, H-skew-Hermitian, H-odd, H-even, H-palindromic$\}$. For $\mathrm{L} \in \mathbb{S}$, consider the set*

$$\mathbb{S}(\lambda, x, \mathrm{L}) := \{\mathrm{K} \in \mathbb{S} : \mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0\}.$$

Then $\mathbb{S}(\lambda, x, \mathrm{L}) \neq \emptyset$ and there exists a unique $\triangle \mathrm{L} \in \mathbb{S}(\lambda, x, \mathrm{L})$ such that

$$\min\{|||\mathrm{K}||| : \mathrm{K} \in \mathbb{S}(\lambda, x, \mathrm{L})\} = |||\triangle \mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L}).$$

Further, each pencil in $\mathbb{S}(\lambda, x, \mathrm{L})$ is of the form $\triangle \mathrm{L} + (I - xx^H)^* \, \mathrm{Z}(I - xx^H)$ for some $\mathrm{Z} \in \mathbb{S}$, where $*$ is either the transpose or the conjugate transpose depending upon the structure defined by $\mathbb{S}$. In other words, we have $\mathbb{S}(\lambda, x, \mathrm{L}) = \triangle \mathrm{L} + (I - xx^H)^* \, \mathbb{S}\, (I - xx^H)$.

We mention that the results obtained above are easily extended to the case of pencils having more general structures. Indeed, let M be a unitary matrix such that $M^T = M$ or $M^T = -M$. Consider the Jordan algebra $\mathbb{J} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^T M = A\}$ and the Lie algebra $\mathbb{L} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^T M = -A\}$ associated with the scalar product $(x, y) \mapsto y^T M x$. Consider a pencil $\mathrm{L}(z) = A + zB$, where A and B are in $\mathbb{J}$ and/or in $\mathbb{L}$. Then the pencil $M\mathrm{L}$ given by $M\mathrm{L}(z) = MA + zMB$ is either T-symmetric, T-skew-symmetric, T-even or T-odd. Hence replacing A, B and r by MA, MB and Mr, respectively, in the above results, we obtain corresponding results for the pencil L.

Similarly, when M is unitary and $M = M^H$ or $M = -M^H$, we consider the Jordan algebra $\mathbb{J} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^H M = A\}$ and the Lie algebra $\mathbb{L} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^H M = -A\}$ associated with the scalar product $(x, y) \mapsto y^H M x$. Now, let $\mathrm{L}(z) = A + zB$ be a pencil where A and B are in $\mathbb{J}$ and/or in $\mathbb{L}$. Then the pencil $M\mathrm{L}(z) = MA + zMB$ is either H-Hermitian, H-skew-Hermitian, H-even or H-odd. Hence replacing A, B and r by MA, MB and Mr, respectively, in the above results, we obtain corresponding results for the pencil L. In particular, when $M := J$, where $J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$, the Jordan algebra $\mathbb{J}$ consists of skew-Hamiltonian matrices and the Lie algebra $\mathbb{L}$ consists of Hamiltonian matrices. So, for example, considering the pencil $\mathrm{L}(z) := A + zB$, where A is Hamiltonian and B is skew-Hamiltonian, we see that the pencil $J\mathrm{L}(z) = JA + zJB$ is H-even. Hence extending the results obtained for H-even pencil to the case of L, we have the following.

TH-789_04612301

Theorem 3.3.10. *Let $\mathbb{S}$ be the space of pencils of the form $\mathrm{L}(z) = A + zB$, where A is Hamiltonian and B is skew-Hamiltonian. For $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, set $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \dfrac{\sqrt{2\|r\|_2^2 - |x^H Jr|^2}}{\|(1, \lambda)\|_2} \leq \sqrt{2}\,\eta(\lambda, x, \mathrm{L}) & \text{if } \lambda \in i\mathbb{R}, \\ \sqrt{|x^H JAx|^2 + |x^H JBx|^2 + \dfrac{2\|r\|_2^2 - 2|x^H Jr|^2}{1 + |\lambda|^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\mathbb{R}. \end{cases}$$

We mention that Remark 3.3.9 remains valid for structured pencils in $\mathbb{S}$ whose coefficient matrices are element of Jordan and/or Lie algebras associated with a scalar product considered above. In such a case the $*$ in $(I - xx^H)^*$ is the adjoint induced by the scalar product that defines the Jordan and Lie algebras.

3.4 Spectral norm and structured backward errors

Considering Frobenius norm on $\mathbb{C}^{n\times n}$, in the previous section, we have obtained structured backward error of an approximate eigenpair. In this section, we derive structured backward errors when $\mathbb{C}^{n\times n}$ is equipped with the spectral norm. Recall that the norm of a pencil $\mathrm{L}(z) = A + zB$ as defined in (3.1) is then given by $|||\mathrm{L}||| := (\|A\|_2^2 + \|B\|_2^2)^{1/2}$. Derivations of structured backward errors of approximate eigenpairs turn out to be much more difficult when $\mathbb{C}^{n\times n}$ is equipped with the spectral norm than in the case when $\mathbb{C}^{n\times n}$ is equipped with the Frobenius norm. We mention that for certain structures (e.g., T-symmetric and T-skew-symmetric) it is indeed possible to use structured mapping Theorems given in chapter 2 (see also [70]) to derive structured backward errors of approximate eigenpairs. However, for most structures (e.g., even, odd, palindromic, Hermitian, skew-Hermitian), the structured mapping Theorems are not of much help for deriving structured backward errors. We overcome this difficulty by employing Davis-Kahan-Weinberger solutions of norm preserving dilation problem for Hilbert space operators. For a more general version of the DKW Theorem, see [24].

We now use the DKW Theorem 1.2.5 with $Z = 0$ and derive structured backward error of an approximate eigenpair. Recall that for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, our standing assumption is that $x^H x = 1$.

Theorem 3.4.1. *Let $\mathbb{S} \in \{T\text{-symmetric, } T\text{-skew-symmetric}\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) := A + zB$. Let $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \dfrac{\|r\|_2}{\|(1, \lambda)\|_2} = \eta(\lambda, x, \mathrm{L})$. Now define*

$$\triangle A := \begin{cases} \frac{1}{1+|\lambda|^2}[\overline{x}r^T + rx^H - (r^T x)\overline{x}x^H - \frac{\overline{x^T r}\;(I-\overline{x}x^T)rr^T(I-xx^H)}{\|r\|_2^2 - |x^T r|^2}], & \text{if } A = A^T, \\ -\frac{1}{1+|\lambda|^2}[\overline{x}r^T - rx^H], & \text{if } A = -A^T. \end{cases}$$

$$\triangle B := \begin{cases} \frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x}r^T + rx^H - (r^T x)\overline{x}x^H - \frac{\overline{x^T r}\;(I-\overline{x}x^T)rr^T(I-xx^H)}{\|r\|_2^2 - |x^T r|^2}], & \text{if } B = B^T, \\ -\frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x}r^T - rx^H], & \text{if } B = -B^T. \end{cases}$$

and consider the pencil $\triangle\mathrm{L}(z) := \triangle A + z\triangle B$. Then $\triangle\mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ and $|||\triangle\mathrm{L}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

TH-789_04612301

Proof: Suppose that L is T-symmetric. Then from the proof of Theorem 3.3.1, we have

$$\triangle A = \overline{Q} \begin{pmatrix} \frac{x^T r}{1+|\lambda|^2} & \frac{1}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2}(Q_1^T r) & A_1 \end{pmatrix} Q^H, \ \triangle B = \overline{Q} \begin{pmatrix} \frac{\overline{\lambda}\, x^T r}{1+|\lambda|^2} & \frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r) & B_1 \end{pmatrix} Q^H,$$

such that $\triangle \mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. Now, for $\mu_{\triangle A} := \frac{\|r\|_2}{1+|\lambda|^2}$ and $\mu_{\triangle B} := \frac{|\lambda|\ \|r\|_2}{1+|\lambda|^2}$, by the DKW Theorem 1.2.5, we have $A_1 = -\dfrac{\overline{x^T r}\ (Q_1^T r)(Q_1^T r)^T}{(1+|\lambda|^2)\ (\|r\|_2^2 - |x^T r|^2)}$ and $B_1 = -\dfrac{\overline{\lambda}\ \overline{x^T r}\ (Q_1^T r)(Q_1^T r)^T}{(1+|\lambda|^2)\ (\|r\|_2^2 - |x^T r|^2)}$. This gives $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = (\|\triangle A\|_2^2 + \|\triangle B\|_2^2)^{1/2} = \dfrac{\|r\|_2}{\|(1,\lambda)\|_2}$. Simplifying expressions for $\triangle A$ and $\triangle B$, we obtain the desired results.

When L is T-skew-symmetric, from the proof of Theorem 3.3.3, we have

$$\triangle A = \overline{Q} \begin{pmatrix} 0 & -\frac{(Q_1^T r)^T}{1+|\lambda|^2} \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix} Q^H, \ \triangle B = \overline{Q} \begin{pmatrix} 0 & -\frac{\overline{\lambda}(Q_1^T r)^T}{1+|\lambda|^2} \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^T r & B_1 \end{pmatrix} Q^H,$$

such that $\triangle \mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. Now, for $\mu_{\triangle A} := \frac{\|r\|_2}{1+|\lambda|^2}$ and $\mu_{\triangle B} := \frac{|\lambda|\ \|r\|_2}{1+|\lambda|^2}$, by the DKW Theorem 1.2.5, we obtain $A_1 = 0 = B_1$. Consequently, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = (\|\triangle A\|_2^2 + \|\triangle B\|_2^2)^{1/2} = \|r\|_2 / \|(1,\lambda)\|_2$. Simplifying the expressions for $\triangle A$ and $\triangle B$, we obtain the desired results. ■

Remark 3.4.2. *If $|x^T r| = \|r\|_2$, then $\|Q_1^T r\|_2 = 0$. In such a case, considering $A_1 = 0 = B_1$ we obtain the desired results.*

Next, we consider T-even and T-odd pencils. Recall that for $z \in \mathbb{C}$, $\mathrm{sign}(z) := \overline{z}/|z|$ when $z \neq 0$ and $\mathrm{sign}(z) := 1$ when $z = 0$.

Theorem 3.4.3. *Let $\mathbb{S} \in \{T\text{-even}, T\text{-odd}\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) := A + zB$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{|x^T A x|^2 + \dfrac{\|r\|_2^2 - |x^T r|^2}{1+|\lambda|^2}} = \dfrac{\sqrt{\|r\|_2^2 + |\lambda|^2 |x^T r|^2}}{\|(1,\ \lambda)\|_2}, & \text{if L is } T\text{-even}, \\ \sqrt{|x^T B x|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{1+|\lambda|^2}} = \dfrac{\sqrt{\|r\|_2^2 + |\lambda|^{-2}\, |x^T r|^2}}{\|(1,\ \lambda)\|_2}, & \text{if L is } T\text{-odd}, \lambda \neq 0, \\ \eta(\lambda, x, \mathrm{L}), & \text{if L is } T\text{-odd}, \lambda = 0. \end{cases}$$

Now, define

$$\triangle A := \begin{cases} -\overline{x} x^T A x x^H + \frac{1}{1+|\lambda|^2}[\overline{x} r^T + r x^H - 2(x^T r)\overline{x} x^H] + \frac{\overline{x^T A x}\ (I - \overline{x} x^T) r r^T (I - x x^H)}{\|r\|_2^2 - |x^T r|^2}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } A = A^T, \\ \frac{1}{1+|\lambda|^2}[r x^H - \overline{x} r^T], \ \text{if } A = -A^T. \end{cases}$$

$$\triangle B := \begin{cases} -\overline{x} x^T B x x^H + \frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x} r^T + r x^H - 2(x^T r)\overline{x} x^H] - \frac{\mathrm{sign}(\lambda)^2\ \overline{x^T B x}\ (I - \overline{x} x^T) r r^T (I - x x^H)}{\|r\|_2^2 - |x^T r|^2}, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } B = B^T. \\ -\frac{\overline{\lambda}}{1+|\lambda|^2}[\overline{x} r^T - r x^H], \ \text{if } B = -B^T, \end{cases}$$

and consider the pencil $\triangle \mathrm{L}(z) := \triangle A + z \triangle B$. Then $\triangle \mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle \mathrm{L}(\lambda)x = 0$ and $\|\triangle \mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

TH-789_04612301

Proof: Suppose that L is T-even. Then from the proof of Theorem 3.3.4, we have

$$\triangle A = \overline{Q}\begin{pmatrix} -x^T A x & \frac{(Q_1^T r)^T}{1+|\lambda|^2} \\ \frac{Q_1^T r}{1+|\lambda|^2} & A_1 \end{pmatrix} Q^H, \ \triangle B = \overline{Q}\begin{pmatrix} 0 & -\frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r) & B_1 \end{pmatrix} Q^H,$$

such that $\triangle \mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. Now, for

$$\mu_{\triangle A} := \sqrt{|x^T Ax|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{(1+|\lambda|^2)^2}} \text{ and } \mu_{\triangle B} := \sqrt{\frac{|\lambda|^2(\|r\|_2^2 - |x^T r|^2)}{(1+|\lambda|^2)^2}}$$

by the DKW Theorem 1.2.5, we have $A_1 = \dfrac{\overline{x^T Ax}}{\|r\|_2^2 - |x^T r|^2}(Q_1^T r)(Q_1^T r)^T$ and $B_1 = 0$. This gives

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \ = \ \sqrt{|x^T Ax|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{1+|\lambda|^2}} = \frac{\sqrt{\|r\|_2^2 + |\lambda|^2 |x^T r|^2}}{\|(1,\ \lambda)\|_2}.$$

Simplifying the expressions for $\triangle A$ and $\triangle B$, we obtain the desired results. When L is T-odd the results follow by interchanging the role of A and B. ■

It follows that for a T-even pencil we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \|(1, \lambda)\|_2 \, \eta(\lambda, x, \mathrm{L})$ whereas for a T-odd pencil we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) \leq \|(1, \lambda^{-1})\|_2 \, \eta(\lambda, x, \mathrm{L})$ when $\lambda \neq 0$.

Theorem 3.4.4. *Let $\mathbb{S} \in \{H\text{-Hermitian}, H\text{-skew-Hermitian}\}$ and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) := A + zB$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \eta(\lambda, x, \mathrm{L}), & \text{if } \ \lambda \in \mathbb{R}, \\ \sqrt{|x^H Ax|^2 + |x^H Bx|^2 + \dfrac{\|r\|_2^2 - |x^H r|^2}{1+|\lambda|^2}}, & \text{if } \lambda \in \mathbb{C} \setminus \mathbb{R} \end{cases}$$

When $\lambda \in \mathbb{R}$, define

$$\triangle A \ := \begin{cases} \frac{1}{1+\lambda^2}[xr^H + rx^H - (r^H x)xx^H - \frac{x^H r \ (I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } A = A^H, \\ \frac{1}{1+\lambda^2}[rx^H - xr^H + (r^H x)xx^H + \frac{r^H x(I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } A = -A^H. \end{cases}$$

$$\triangle B \ := \begin{cases} \frac{\lambda}{1+\lambda^2}[xr^H + rx^H - (r^H x)xx^H - \frac{x^H r \ (I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } B = B^H, \\ \frac{\lambda}{1+\lambda^2}[rx^H - xr^H + (r^H x)xx^H + \frac{r^H x \ (I - xx^H)rr^H(I - xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } B = -B^H. \end{cases}$$

TH-789_04612301

When $\lambda \in \mathbb{C}\setminus\mathbb{R}$, *define*

$$\triangle A \;:=\; \begin{cases} -xx^HAxx^H + \frac{1}{1+|\lambda|^2}[xr^H(I-xx^H)+(I-xx^H)rx^H] + \frac{x^HAx\;(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2-|x^Hr|^2}, \\ \hfill \text{if } A=A^H, \\ -xx^HAxx^H + \frac{1}{1+|\lambda|^2}[(I-xx^H)rx^H - xr^H(I-xx^H)] + \frac{x^HAx(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2-|x^Hr|^2}, \\ \hfill \text{if } A=-A^H. \end{cases}$$

$$\triangle B \;:=\; \begin{cases} -xx^HBxx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I-xx^H)+\overline{\lambda}(I-xx^H)rx^H] + \frac{x^HBx\;(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2-|x^Hr|^2}, \\ \hfill \text{if } B=B^H, \\ -xx^HBxx^H - \frac{\lambda}{1+|\lambda|^2}xr^H(I-xx^H) + \frac{\overline{\lambda}}{1+|\lambda|^2}(I-xx^H)rx^H + \frac{x^HBx(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2-|x^Hr|^2}, \\ \hfill \text{if } B=-B^H. \end{cases}$$

Consider $\triangle\mathrm{L}(z) := \triangle A + z\triangle B$. *Then* $\triangle\mathrm{L}\in\mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $|||\triangle\mathrm{L}||| = \eta^{\mathbb{S}}(\lambda,x,\mathrm{L})$.

Proof: First, suppose that L is H-Hermitian. Assume that $\lambda\in\mathbb{R}$. Then $x^Hr\in\mathbb{R}$. Now from the proof of Theorem 3.3.6, we have

$$\triangle A = Q\begin{pmatrix} \frac{1}{1+\lambda^2}x^Hr & \frac{1}{1+\lambda^2}(Q_1^Hr)^H \\ \frac{1}{1+\lambda^2}Q_1^Hr & A_1 \end{pmatrix}Q^H \text{ and } \triangle B = Q\begin{pmatrix} \frac{\lambda}{1+\lambda^2}x^Hr & \frac{\lambda}{1+\lambda^2}(Q_1^Hr)^H \\ \frac{\lambda}{1+\lambda^2}Q_1^Hr & B_1 \end{pmatrix}Q^H$$

such that $\triangle\mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. For $\mu_{\triangle A} := \frac{\|r\|_2}{1+\lambda^2}$ and $\mu_{\triangle B} := \frac{|\lambda|\,\|r\|_2}{1+\lambda^2}$ by the DKW Theorem 1.2.5, we have $A_1 = -\dfrac{x^Hr\;(Q_1^Hr)(Q_1^Hr)^H}{(1+\lambda^2)\;(\|r\|_2^2-|x^Hr|^2)}$, $B_1 = -\dfrac{\lambda\, x^Hr\;(Q_1^Hr)(Q_1^Hr)^H}{(1+\lambda^2)\;(\|r\|_2^2-|x^Hr|^2)}$. This gives

$$\eta^{\mathbb{S}}(\lambda,x,\mathrm{L}) = (\|\triangle A\|_2^2 + \|\triangle B\|_2^2)^{1/2} = \frac{\|r\|_2}{\|(1,\lambda)\|_2}.$$

Now simplifying the expressions for $\triangle A$ and $\triangle B$, we obtain the desired results.

Next, suppose that $\lambda\in\mathbb{C}\setminus\mathbb{R}$. Then again from the proof of Theorem 3.3.6, we have

$$\triangle A = Q\begin{pmatrix} -x^HAx & \frac{1}{1+|\lambda|^2}(Q_1^Hr)^H \\ \frac{1}{1+|\lambda|^2}Q_1^Hr & A_1 \end{pmatrix}Q^H,\; \triangle B := Q\begin{pmatrix} -x^HBx & \frac{\lambda}{1+|\lambda|^2}(Q_1^Hr)^H \\ \frac{\overline{\lambda}}{1+|\lambda|^2}Q_1^Hr & B_1 \end{pmatrix}Q^H.$$

For, $\mu_{\triangle A} := \sqrt{|x^HAx|^2 + \frac{\|r\|_2^2-|x^Hr|^2}{(1+|\lambda|^2)^2}}$, $\mu_{\triangle B} := \sqrt{|x^HBx|^2 + \frac{|\lambda|^2(\|r\|_2^2-|x^Hr|^2)}{(1+|\lambda|^2)^2}}$, by the DKW Theorem 1.2.5, we have

$$A_1 = \frac{x^HAx}{\|r\|_2^2-|x^Hr|^2}(Q_1^Hr)(Q_1^Hr)^H \text{ and } B_1 = \frac{x^HBx}{\|r\|_2^2-|x^Tr|^2}(Q_1^Hr)(Q_1^Hr)^H.$$

Hence we have $\eta^{\mathbb{S}}(\lambda,x,\mathrm{L}) = \sqrt{|x^HAx|^2 + |x^HBx|^2 + \dfrac{\|r\|_2^2-|x^Hr|^2}{1+|\lambda|^2}}$. Now, simplifying the expressions for $\triangle A$ and $\triangle B$, we obtain the desired results. The proof is similar for the case when L is H-skew-Hermitian. ■

We mention that when $Q_1^Hr = 0$, the desired results follow by considering $A_1 = 0 = B_1$.

Theorem 3.4.5. *Let* $\mathbb{S}\in\{H\text{-even}, H\text{-odd}\}$ *and* $\mathrm{L}\in\mathbb{S}$ *be given by* $\mathrm{L}(z) := A + zB$. *Let*

TH-789_04612301

$(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ *and* $r := -\mathrm{L}(\lambda)x$. *Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \eta(\lambda, x, \mathrm{L}), & \text{if } \lambda \in i\mathbb{R}, \\ \sqrt{|x^H Ax|^2 + |x^H Bx|^2 + \dfrac{\|r\|_2^2 - |x^H r|^2}{1+|\lambda|^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\mathbb{R}. \end{cases}$$

When $\lambda \in i\mathbb{R}$, *define*

$$\triangle A := \begin{cases} \frac{1}{1+|\lambda|^2}[xr^H + rx^H - (r^H x)xx^H - \frac{x^H r\ (I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } A = A^H, \\ \frac{1}{1+|\lambda|^2}[rx^H - xr^H + (r^H x)xx^H + \frac{r^H x(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } A = -A^H. \end{cases}$$

$$\triangle B := \begin{cases} \frac{1}{1+|\lambda|^2}[\overline{\lambda} rx^H + \lambda xr^H - \lambda(r^H x)xx^H + \frac{\lambda x^H r\ (I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \text{if } B = B^H, \\ \frac{1}{1+|\lambda|^2}[\overline{\lambda} xx^H rx^H - \lambda xr^H(I-xx^H) + \overline{\lambda}(I-xx^H)rx^H + \frac{\lambda r^H x(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}], & \\ \hfill \text{if } B = -B^H. \end{cases}$$

When $\lambda \in \mathbb{C} \setminus i\mathbb{R}$, *define*

$$\triangle A := \begin{cases} -xx^H Axx^H + \frac{1}{1+|\lambda|^2}[xr^H(I-xx^H) + (I-xx^H)rx^H] + \frac{x^H Ax\ (I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}, \\ \hfill \text{if } A = A^H, \\ -xx^H Axx^H + \frac{1}{1+|\lambda|^2}[(I-xx^H)rx^H - xr^H(I-xx^H)] + \frac{x^H Ax(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}, \\ \hfill \text{if } A = -A^H. \end{cases}$$

$$\triangle B := \begin{cases} -xx^H Bxx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I-xx^H) + \overline{\lambda}(I-xx^H)rx^H] + \frac{x^H Bx\ (I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}, \\ \hfill \text{if } B = B^H, \\ -xx^H Bxx^H - \frac{\lambda xr^H(I-xx^H)}{1+|\lambda|^2} + \frac{\overline{\lambda}(I-xx^H)rx^H}{1+|\lambda|^2} + \frac{x^H Bx(I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2 - |x^H r|^2}, \\ \hfill \text{if } B = -B^H. \end{cases}$$

Consider $\triangle\mathrm{L}(z) := \triangle A + z\triangle B$. *Then* $\triangle\mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: First, suppose that L is H-even. Next, assume that $\lambda \in i\mathbb{R}$. Then it follows that $x^H r \in \mathbb{R}$. Now from the proof of Theorem 3.3.7, we have

$$\triangle A = Q \begin{pmatrix} \frac{1}{1+|\lambda|^2} x^H r & \frac{1}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \ \triangle B := Q \begin{pmatrix} \frac{\overline{\lambda}}{1+|\lambda|^2} x^H r & -\frac{\lambda}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r & B_1 \end{pmatrix} Q^H$$

such that $\triangle\mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. For $\mu_{\triangle A} := \frac{\|r\|_2}{1+|\lambda|^2}$, $\mu_{\triangle B} := \frac{|\lambda|\ \|r\|_2}{1+|\lambda|^2}$, by the DKW Theorem 1.2.5 we have

$$A_1 = -\frac{x^H r\ (Q_1^H r)(Q_1^H r)^H}{(1+|\lambda|^2)\ (\|r\|_2^2 - |x^H r|^2)} \text{ and } B_1 = \frac{\lambda\ x^H r\ (Q_1^H r)(Q_1^H r)^H}{(1+|\lambda|^2)\ (\|r\|_2^2 - |x^H r|^2)}.$$

This gives $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = (\|\triangle A\|_2^2 + \|\triangle B\|_2^2)^{1/2} = \dfrac{\|r\|_2}{\|(1,\lambda)\|_2}$. Simplifying expressions for $\triangle A$ and $\triangle B$, we obtain the desired result.

Now suppose that $\lambda \in \mathbb{C} \setminus i\mathbb{R}$. The again from the proof of Theorem 3.3.7, we have

$$\triangle A = Q \begin{pmatrix} -x^H Ax & \frac{1}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2} Q_1^H r & A_1 \end{pmatrix} Q^H, \ \triangle B := Q \begin{pmatrix} -x^H Bx & -\frac{\lambda}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{\overline{\lambda}}{1+|\lambda|^2} Q_1^H r & B_1 \end{pmatrix} Q^H.$$

TH-789_04612301

In this case we have, $\mu_{\triangle A} = \sqrt{|x^HAx|^2 + \frac{\|r\|_2^2 - |x^Hr|^2}{(1+|\lambda|^2)^2}}$, $\mu_{\triangle B} = \sqrt{|x^HBx|^2 + \frac{|\lambda|^2(\|r\|_2^2 - |x^Hr|^2)}{(1+|\lambda|^2)^2}}$. By the DKW Theorem 1.2.5, we have

$$A_1 = \frac{x^HAx}{\|r\|_2^2 - |x^Hr|^2}(Q_1^Hr)(Q_1^Hr)^H \text{ and } B_1 = \frac{x^HBx}{\|r\|_2^2 - |x^Tr|^2}(Q_1^Hr)(Q_1^Hr)^H.$$

Consequently, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{|x^HAx|^2 + |x^HBx|^2 + \frac{\|r\|_2^2 - |x^Hr|^2}{1+|\lambda|^2}}$. Now, simplifying the expressions for $\triangle A$ and $\triangle B$, we obtain the desired results.

When L is H-odd, the desired results follow by interchanging the role of A and B.■

As before, the above results are easily extended to the case of general structured pencils where the coefficient matrices are elements of Jordan and/or Lie algebras. In particular, for the pencil $\mathrm{L}(z) := A + zB$, where A is Hamiltonian and B is skew-Hamiltonian, we have the following result.

Theorem 3.4.6. *Let $\mathbb{S}$ be the space of pencils of the form $\mathrm{L}(z) = A + zB$, where A is Hamiltonian and B is skew-Hamiltonian. Let $\mathrm{L} \in \mathbb{S}$ and $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Set $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \eta(\lambda, x, \mathrm{L}), & \text{if } \lambda \in i\mathbb{R} \\ \sqrt{|x^HJAx|^2 + |x^HJBx|^2 + \frac{\|r\|_2^2 - |x^HJr|^2}{1+|\lambda|^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\mathbb{R}. \end{cases}$$

Now we consider palindromic pencils.

Theorem 3.4.7. *Let $\mathbb{S}$ be the space of T-palindromic pencils and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) := A + zA^T$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\sqrt{|x^TAx|^2 + \frac{|\lambda|^2\,(\|r\|_2^2 - |x^Tr|^2)}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| > 1, \\ \sqrt{2}\sqrt{|x^TAx|^2 + \frac{\|r\|_2^2 - |x^Tr|^2}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| \leq 1 \text{ and } \lambda \neq \pm 1, \\ \eta(\lambda, x, \mathrm{L}), & \text{if } \lambda = \pm 1. \end{cases}$$

Now define

$$\triangle A := \begin{cases} -\overline{x}x^TAxx^H + \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H] + \frac{\overline{\lambda}\ \overline{x^TAx}\ (I-\overline{x}x^T)rr^T(I-xx^H)}{|\lambda|^2\ (\|r\|_2^2 - |x^Tr|^2)}, \\ \qquad \text{if } |\lambda| > 1, \\ -\overline{x}x^TAxx^H + \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H] + \frac{\overline{\lambda}\ \overline{x^TAx}\ (I-\overline{x}x^T)rr^T(I-xx^H)}{\|r\|_2^2 - |x^Tr|^2}, \\ \qquad \text{if } |\lambda| \leq 1 \text{ and } \lambda \neq -1, \\ \frac{1}{1+|\lambda|^2}[\overline{\lambda}\overline{x}r^T(\mathrm{I} - xx^H) + (\mathrm{I} - \overline{x}x^T)rx^H], \quad \text{if } \lambda = -1. \end{cases}$$

Consider the pencil $\triangle\mathrm{L}(z) := \triangle A + z(\triangle A)^T$. Then $\triangle\mathrm{L} \in \mathbb{S}$, $\mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ and $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

TH-789_04612301

Proof: Suppose that $\lambda \neq -1$. Then from the proof of Theorem 3.3.5, we have

$$\triangle A = \overline{Q} \begin{pmatrix} -x^T A x & \frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix} Q^H$$

such that $\triangle \mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. Now for

$$\mu_{\triangle A} := \begin{cases} \sqrt{|x^T A x|^2 + \frac{|\lambda|^2(\|r\|_2^2 - |x^T r|^2)}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| > 1, \\ \sqrt{|x^T A x|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| \leq 1, \end{cases}$$

by the DKW Theorem 1.2.5, we have

$$A_1 = \begin{cases} \frac{\overline{\lambda}\ \overline{x^T A x}}{|\lambda|^2\ (\|r\|_2^2 - |x^T r|^2)} Q_1^T r (Q_1^T r)^T, & \text{if } |\lambda| > 1, \\ \frac{\overline{\lambda}\ \overline{x^T A x}}{\|r\|_2^2 - |x^T r|^2} Q_1^T r (Q_1^T r)^T, & \text{if } |\lambda| \leq 1. \end{cases}$$

Consequently, we have

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\ \sqrt{|x^T A x|^2 + \frac{|\lambda|^2\ (\|r\|_2^2 - |x^T r|^2)}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| > 1, \\ \sqrt{2}\ \sqrt{|x^T A x|^2 + \frac{\|r\|_2^2 - |x^T r|^2}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

Now simplifying the expression for $\triangle A$, we obtain the desired results.

Next, suppose that $\lambda = -1$. Then again from the proof of Theorem 3.3.5, we have

$$\triangle A = \overline{Q} \begin{pmatrix} 0 & \frac{\overline{\lambda}}{1+|\lambda|^2}(Q_1^T r)^T \\ \frac{1}{1+|\lambda|^2} Q_1^T r & A_1 \end{pmatrix} Q^H. \text{ For } \mu_{\triangle A} := \frac{\|r\|_2}{1+|\lambda|^2},$$

by the DKW Theorem 1.2.5, we have $A_1 = 0$. Hence $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \frac{1}{\sqrt{2}}\|r\|_2$. Simplifying the expression for $\triangle A$, we obtain the desired result. ■

For H-palindromic pencils we have the following.

Theorem 3.4.8. *Let $\mathbb{S}$ be the space of H-palindromic pencils and $\mathrm{L} \in \mathbb{S}$ be given by $\mathrm{L}(z) := A + zA^H$. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $r := -\mathrm{L}(\lambda)x$. Then we have*

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\ \sqrt{|x^H A x|^2 + \dfrac{|\lambda|^2\ (\|r\|_2^2 - |x^H r|^2)}{(1+|\lambda|^2)^2}}, & \textit{if } |\lambda| > 1, \\ \sqrt{2}\ \sqrt{|x^H A x|^2 + \dfrac{\|r\|_2^2 - |x^H r|^2}{(1+|\lambda|^2)^2}}, & \textit{if } |\lambda| < 1, \\ \eta(\lambda, x, \mathrm{L}), & \textit{if } |\lambda| = 1. \end{cases}$$

TH-789_04612301

Now define

$$\triangle A := \begin{cases} -xx^H Axx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I - xx^H) + (I - xx^H)rx^H] + \frac{\overline{x^H Ax}\ (I-xx^H)rr^H(I-xx^H)}{\overline{\lambda}\,(\|r\|_2^2-|x^H r|^2)}, \\ \hfill \text{if } |\lambda| > 1, \\ -xx^H Axx^H + \frac{1}{1+|\lambda|^2}[\lambda xr^H(I - xx^H) + (I - xx^H)rx^H] + \frac{\lambda\ \overline{x^H Ax}\ (I-xx^H)rr^H(I-xx^H)}{\|r\|_2^2-|x^H r|^2}, \\ \hfill \text{if } |\lambda| < 1, \\ \frac{1}{1+|\lambda|^2}[rx^H + \lambda xr^H(\mathrm{I} - xx^H) - \frac{x^H r\ (I-xx^H)rr^H(I-xx^H)}{(\|r\|_2^2-|x^H r|^2)}], \ \text{if } |\lambda| = 1, \end{cases}$$

and consider the pencil $\triangle\mathrm{L}(z) := \triangle A + z(\triangle A)^H$*. Then* $\triangle\mathrm{L} \in \mathbb{S},\ \mathrm{L}(\lambda)x + \triangle\mathrm{L}(\lambda)x = 0$ *and* $\|\triangle\mathrm{L}\| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$.

Proof: First, suppose that $|\lambda| \neq 1$. Then from the proof of Theorem 3.3.8, we have

$$\triangle A = Q \begin{pmatrix} -x^H Ax & \frac{\lambda}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2}Q_1^H r & A_1 \end{pmatrix} Q^H$$

such that $\triangle\mathrm{L}(\lambda)x + \mathrm{L}(\lambda)x = 0$. In this case, we have

$$\mu_{\triangle A} = \begin{cases} \sqrt{|x^H Ax|^2 + \frac{|\lambda|^2(\|r\|_2^2-|x^H r|^2)}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| > 1, \\ \sqrt{|x^H Ax|^2 + \frac{\|r\|_2^2-|x^H r|^2}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| < 1. \end{cases}$$

Hence by the DKW Theorem 1.2.5, we have

$$A_1 = \begin{cases} \frac{\lambda\ \overline{x^H Ax}}{|\lambda|^2\ (\|r\|_2^2-|x^H r|^2)} Q_1^H r(Q_1^H r)^H, & \text{if } |\lambda| > 1, \\ \frac{\lambda\ \overline{x^H Ax}}{\|r\|_2^2-|x^H r|^2} Q_1^H r(Q_1^H r)^H, & \text{if } |\lambda| < 1. \end{cases}$$

This gives

$$\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \begin{cases} \sqrt{2}\ \sqrt{|x^H Ax|^2 + \frac{|\lambda|^2\ (\|r\|_2^2-|x^H r|^2)}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| > 1, \\ \sqrt{2}\ \sqrt{|x^H Ax|^2 + \frac{\|r\|_2^2-|x^H r|^2}{(1+|\lambda|^2)^2}}, & \text{if } |\lambda| < 1. \end{cases}$$

Simplifying the expression for $\triangle A$, we obtain the desired result.

When $|\lambda| = 1$, again from the proof of Theorem 3.3.8, we have

$$\triangle A = Q \begin{pmatrix} \frac{x^H r}{1+|\lambda|^2} & \frac{\lambda}{1+|\lambda|^2}(Q_1^H r)^H \\ \frac{1}{1+|\lambda|^2}Q_1^H r & A_1 \end{pmatrix} Q^H$$

Now, we have $\mu_{\triangle A} = \dfrac{\|r\|_2}{1+|\lambda|^2}$. Hence by the DKW Theorem 1.2.5, we have

$$A_1 = -\frac{x^H r\ (I - xx^H)rr^H(I - xx^H)}{(1+|\lambda|^2)(\|r\|_2^2 - |x^H r|^2)}.$$

Consequently, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \dfrac{\|r\|_2}{\sqrt{2}}$. Simplifying the expression for $\triangle A$, we obtain the desired result. ■

Remark 3.4.9. *Let* $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$ *with* $x^H x = 1$ *and* $\mathbb{S} \in \{$*T-symmetric, T-skew-symmetric,*

TH-789_04612301

T-odd, T-even, T-palindromic, H-Hermitian, H-skew-Hermitian, H-odd, H-even, H-palindromic}. For $\mathrm{L} \in \mathbb{S}$, consider the set

$$\mathbb{S}(\lambda, x, \mathrm{L}) := \{\mathrm{K} \in \mathbb{S} : \mathrm{L}(\lambda)x + \mathrm{K}(\lambda)x = 0\}.$$

Then $\mathbb{S}(\lambda, x, \mathrm{L}) \neq \emptyset$ and $\min\{|\!|\!|\mathrm{K}|\!|\!| : \mathrm{K} \in \mathbb{S}(\lambda, x, \mathrm{L})\} = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$. Further,

$$\mathbb{S}_{\mathrm{opt}}(\lambda, x, \mathrm{L}) := \{\triangle\mathrm{L} \in \mathbb{S}(x, \lambda, \mathrm{L}) : |\!|\!|\triangle\mathrm{L}|\!|\!| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{L})\}$$

is an infinite set and is characterized by the DKW Theorem 1.2.5 by taking into account the nonzero contractions. Let $\triangle\mathrm{L} \in \mathbb{S}_{\mathrm{opt}}(\lambda, x, \mathrm{L})$. Then each pencil in $\mathbb{S}(\lambda, x, \mathrm{L})$ is of the form $\triangle\mathrm{L} + (I - xx^H)^* \, \mathrm{Z}(I - xx^H)$ for some $\mathrm{Z} \in \mathbb{S}$, where $*$ is either the transpose or the conjugate transpose depending upon the structure defined by $\mathbb{S}$. In other words, we have

$$\mathbb{S}(\lambda, x, \mathrm{L}) = \triangle\mathrm{L} + (I - xx^H)^* \, \mathbb{S} \, (I - xx^H).$$

Needless to mention that Remark 3.4.9 remains valid for structured pencils in $\mathbb{S}$ whose coefficient matrices are element of Jordan and/or Lie algebras associated with a scalar product considered in the previous section. In such a case the $*$ in $(I - xx^H)^*$ is the adjoint induced by the scalar product that defines the Jordan and Lie algebras.

We now illustrate various structured and unstructured backward errors by numerical examples. We use MATLAB.7.0 for our computation. We generate A and B as follows:

```
>> randn('state',15), A = randn(50)+ i*randn(50); A = A ± A*;
>> randn('state',25), B = randn(50)+i*randn(50); B = B ± B*;
```

For T-palindromic/H-palindromic pencils, we generate A and B by

```
>> randn('state',15), A = randn(50)+ i*randn(50); B = A*;
```

Here $A^* = A^T$ or $A^* = A^H$. Finally, we compute (λ, x) by

```
>> [V,D] = eig(A,B); λ = -D(2,2); x = V(:,2)/norm(V(:,2));
```

We denote by $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{L})$ and $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{L})$ the backward error $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L})$ when $\mathbb{C}^{n\times n}$ is equipped with the Frobenius norm and the spectral norm, respectively. Note that $\eta(\lambda, x, \mathrm{L})$ is the same for the spectral and the Frobenius norms. Table 3.2 gives the computed result.

Note that structured backward errors are bigger than or equal to unstructured backward errors but they are marginally so. On the other hand, structured condition numbers are less than or equal to unstructured condition numbers [16, 54]. Consequently, structured backward errors when combined with structured condition numbers provide almost the same approximate upper bounds on the errors in the computed eigenelements as do their unstructured counterparts. We mention that the MATLAB `eig` command does not ensure spectral symmetry in the computed eigenvalues.

3.5 Structured pseudospectra of structured pencils

Let $\mathrm{L} \in \mathbb{P}_1(\mathbb{C}^{n\times n})$ be a regular pencil. For $\lambda \in \mathbb{C}$, the backward error of λ as an approximate eigenvalue of L is given by $\eta(\lambda, \mathrm{L}) := \min\{\eta(\lambda, x, \mathrm{L}) : x \in \mathbb{C}^n \text{ and } \|x\|_2 = 1\}$. Since

TH-789_04612301

$\mathbb{S}$	$\eta(\lambda, x, \mathrm{L})$	$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{L})$	$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{L})$
T-symm	1.387705737323579e-014	1.959539856593202e-014	1.387705737323579e-014
T-skew-symm	1.796046101865378e-014	2.539992755905347e-014	1.796046101865378e-014
T-even	2.219610496439476e-014	3.211055813711074e-014	2.324926535413804e-014
T-odd	1.559070464273151e-014	2.204626223816091e-014	1.559075824083717e-014
T-palindromic	1.068704043320177e-014	1.512088705618463e-014	1.487010794022381e-014
H-Herm	2.076731533185186e-014	2.947235222707197e-014	2.106896507205170e-014
H-skew-Herm	1.714743310005108e-014	2.489567700503872e-014	1.811820338752170e-014
H-even	1.590165856939442e-014	2.299115486213681e-014	1.663718384482337e-014
H-odd	2.343032834027323e-014	3.472518481940936e-014	2.566511276851151e-014
H-palindromic	9.161344100487524e-015	1.298942035829892e-014	1.296310627878570e-014

Table 3.2: Numerical computation for structured backward error for structured pencils.

$\eta(\lambda, x, \mathrm{L}) = \|\mathrm{L}(\lambda)x\|_2/\|(1,\lambda)\|_2$, it follows that for the spectral norm as well as for the Frobenius norm on $\mathbb{C}^{n\times n}$, we have $\eta(\lambda, \mathrm{L}) := \sigma_{\min}(\mathrm{L}(\lambda))/\|(1,\lambda)\|_2$. Similarly, we define structured backward error of an approximate eigenvalue λ of $\mathrm{L} \in \mathbb{S}$ by

$$\eta^{\mathbb{S}}(\lambda, \mathrm{L}) := \min\{\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) : x \in \mathbb{C}^n \text{ and } \|x\|_2 = 1\}.$$

Note that backward errors of approximate eigenvalues and pseudospectra of a pencil are closely related. For $\epsilon > 0$, the unstructured ϵ-pseudospectrum of L, denoted by $\sigma_\epsilon(\mathrm{L})$, is given by [3]

$$\sigma_\epsilon(\mathrm{L}) = \bigcup_{\|\triangle\mathrm{L}\|\leq\epsilon} \{\sigma(\mathrm{L} + \triangle\mathrm{L}) : \triangle\mathrm{L} \in \mathbb{P}_1(\mathbb{C}^{n\times n})\}.$$

See [2, 3] for more on pseudospectra of matrix pencils. Obviously, we have $\sigma_\epsilon(\mathrm{L}) = \{z \in \mathbb{C} : \eta(z, \mathrm{L}) \leq \epsilon\}$, assuming, for simplicity, that $\infty \notin \sigma_\epsilon(\mathrm{L})$. For the sake of simplicity, for rest of this section, we make an implicit assumption that $\infty \notin \sigma_\epsilon(\mathrm{L})$. We observe the following.

- Since $\eta(\lambda, \mathrm{L})$ is the same for the spectral norm and the Frobenius norm on $\mathbb{C}^{n\times n}$, it follows that $\sigma_\epsilon(\mathrm{L})$ is the same for the spectral and the Frobenius norms.

Similarly, when $\mathrm{L} \in \mathbb{S}$, we define the structured ϵ-pseudospectrum of L, denoted by $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L})$, by

$$\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) := \bigcup_{\|\triangle\mathrm{L}\|\leq\epsilon} \{\sigma(\mathrm{L} + \triangle\mathrm{L}) : \triangle\mathrm{L} \in \mathbb{S}\}.$$

Then it follows that $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \{z \in \mathbb{C} : \eta^{\mathbb{S}}(\lambda, \mathrm{L}) \leq \epsilon\}$.

Theorem 3.5.1. *Let $\mathbb{S} \in \{T$-symmetric, T-skew-symmetric$\}$ and $\mathrm{L} \in \mathbb{S}$. Let $\lambda \in \mathbb{C}$. Then for the spectral norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \eta(\lambda, \mathrm{L})$ and $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \sigma_\epsilon(\mathrm{L})$. For the Frobenius norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \sqrt{2}\,\eta(\lambda, \mathrm{L})$ and $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \sigma_{\epsilon/\sqrt{2}}(\mathrm{L})$ when L is T-skew-symmetric, and $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \eta(\lambda, \mathrm{L})$ and $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) = \sigma_\epsilon(\mathrm{L})$ when L is T-symmetric.*

Proof: For the spectral norm, by Theorem 3.4.1, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \eta(\lambda, x, \mathrm{L})$ for all x. Consequently, we have $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \eta(\lambda, \mathrm{L})$. Hence the result follows.

For the Frobenius norm, the result follows from Theorem 3.3.3 when L is T-skew-symmetric. So, suppose that L is T-symmetric. Then $\mathrm{L}(\lambda) \in \mathbb{C}^{n\times n}$ is symmetric. Consider the Takagi

TH-789_04612301

factorization $L(\lambda) = U\Sigma U^T$, where U is unitary and Σ is a diagonal matrix containing singular values of $L(\lambda)$ (appear in descending order). Set $s := \Sigma(n,n)$ and $u := U(:,n)$. Then we have $L(\lambda)\overline{u} = su$. Now define $\triangle A := -\frac{s\,uu^T}{1+|\lambda|^2}, \triangle B := -\frac{\overline{\lambda}\,s\,uu^T}{1+|\lambda|^2}$ and consider the pencil $\triangle L(z) = \triangle A + z\triangle B$. Then $\triangle L$ is T-symmetric and $L(\lambda)\overline{u} + \triangle L(\lambda)\overline{u} = 0$. Notice that, for the spectral norm and the Frobenius norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) \leq \|\triangle L\| = s/\|(1,\lambda)\|_2 = \eta(\lambda, L)$ and hence $\sigma_\epsilon(L) = \sigma_\epsilon^{\mathbb{S}}(L)$. This completes the proof. ■

When L is T-symmetric, the above proof shows how to construct a T-symmetric pencil $\triangle L$ such that $\lambda \in \sigma(L + \triangle L)$ and $\|\triangle L\| = \eta^{\mathbb{S}}(\lambda, L)$. When L is T-skew-symmetric, using Takagi factorization of the complex skew-symmetric matrix $L(\lambda)$, one can construct a T-skew-symmetric pencil $\triangle L$ such that $\lambda \in \sigma(L + \triangle L)$ and $\|\triangle L\| = \eta^{\mathbb{S}}(\lambda, L)$. Indeed, consider the Takagi factorization $L(\lambda) = U\mathrm{diag}(d_1, \cdots, d_m)U^T$, where U is unitary, $d_j := \begin{bmatrix} 0 & s_j \\ -s_j & 0 \end{bmatrix}$, $s_j \in \mathbb{C}$ is nonzero and $|s_j|$ are singular values of $L(\lambda)$. Here the blocks d_j appear in descending order of magnitude of $|s_j|$. Note that $L(\lambda)\overline{U} = U\mathrm{diag}(d_1, \cdots, d_m)$. Let $u := U(:, n-1:n)$. Then $L(\lambda)\overline{u} = ud_m = ud_m u^T\overline{u}$. Now define

$$\triangle A := -\frac{ud_m u^T}{1+|\lambda|^2}, \ \ \triangle B := -\frac{\overline{\lambda}\, ud_m u^T}{1+|\lambda|^2}$$

and consider $\triangle L(z) := \triangle A + z\triangle B$. Then $\triangle L$ is T-skew-symmetric and $L(\lambda)\overline{u} + \triangle L(\lambda)\overline{u} = 0$. For the spectral norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) = \|\triangle L\| = \sigma_{\min}(L(\lambda))/\|(1,\lambda)\|_2 = \eta(\lambda, L)$ and for the Frobenius norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) = \|\triangle L\| = \sqrt{2}\,\sigma_{\min}(L(\lambda))/\|(1,\lambda)\|_2 = \sqrt{2}\,\eta(\lambda, L)$.

We denote the unit circle in $\mathbb{C}$ by $\mathbb{T}$, that is, $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. Then for T-even and T-odd pencils we have the following.

Theorem 3.5.2. *Let $\mathbb{S} \in \{T\text{-even}, T\text{-odd}\}$ and $L \in \mathbb{S}$. Let $\lambda \in \mathbb{T}$. Then for the Frobenius norm on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) = \sqrt{2}\,\eta(\lambda, L)$ and $\sigma_\epsilon^{\mathbb{S}}(L) \cap \mathbb{T} = \sigma_{\epsilon/\sqrt{2}}(L) \cap \mathbb{T}$.*

Proof: Let $\lambda \in \mathbb{T}$. Then by Theorem 3.3.4, we have $\eta^{\mathbb{S}}(\lambda, x, L) = \frac{\sqrt{2}\,\|L(\lambda)x\|_2}{\|(1,\lambda)\|_2}$ for all x such that $\|x\|_2 = 1$. Hence taking minimum over $\|x\|_2 = 1$, we obtain the desired results. ■

Theorem 3.5.3. *Let $\mathbb{S} \in \{H\text{-Hermitian}, H\text{-skew-Hermitian}\}$ and $L \in \mathbb{S}$. Let $\lambda \in \mathbb{R}$. Then for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) = \eta(\lambda, L)$ and $\sigma_\epsilon^{\mathbb{S}}(L) \cap \mathbb{R} = \sigma_\epsilon(L) \cap \mathbb{R}$. Also when $\lambda = \pm i$, for the Frobenius norm, we have $\eta^{\mathbb{S}}(\lambda, L) = \sqrt{2}\,\eta(\lambda, L)$.*

Proof: Note that $L(\lambda)$ is either Hermitian or skew-Hermitian. Let (μ, u) be an eigenpair of the matrix $L(\lambda)$ such that $|\mu| = \sigma_{\min}(L(\lambda))$ and $u^H u = 1$. Then $L(\lambda)u = \mu u$. Define

$$\triangle A := -\frac{\mu\, uu^H}{1+|\lambda|^2}, \triangle B := -\frac{\overline{\lambda}\,\mu\, uu^H}{1+|\lambda|^2}$$

and consider the pencil $\triangle L(z) = \triangle A + z\triangle B$. Then $\triangle L \in \mathbb{S}$ and $\lambda \in \Lambda_m(L + \triangle L)$. Further, for the spectral and the Frobenius norms, we have $\|\triangle L\| = \sigma_{\min}(L(\lambda))/\|(1,\lambda)\|_2$. Hence the result follows. Finally, when $\lambda = \pm i$, the result follows from Theorem 3.3.6. ■

Theorem 3.5.4. *Let $\mathbb{S} \in \{H\text{-even}, H\text{-odd}\}$ and $L \in \mathbb{S}$. Let $\lambda \in i\mathbb{R}$. Then for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, L) = \eta(\lambda, L)$ and $\sigma_\epsilon^{\mathbb{S}}(L) \cap i\mathbb{R} = \sigma_\epsilon(L) \cap i\mathbb{R}$. Also*

TH-789_04612301

when $\lambda = \pm 1$, *for the Frobenius norm, we have* $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \sqrt{2}\,\eta(\lambda, \mathrm{L})$.

Proof: Note for $\lambda \in i\mathbb{R}$, the matrix $\mathrm{L}(\lambda)$ is again either is Hermitian or skew-Hermitian. Hence the result follows from the proof of Theorem 3.5.3. When $\lambda = \pm 1$, the result follows from Theorem 3.3.7. ■

We mention that the above results are easily extended to the case of general structured pencils where the coefficients matrices are elements of Jordan and/or Lie algebras.

Finally, for T-palindromic and H-palindromic pencils we have the following result.

Theorem 3.5.5. *Let* $\mathbb{S}$ *be the space of* T*-palindromic pencils and* $\mathrm{L} \in \mathbb{S}$. *Let* $\lambda \in i\mathbb{R}$. *Then for the Frobenius norm on* $\mathbb{C}^{n\times n}$, $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \sqrt{2}\,\eta(\lambda, \mathrm{L})$ *and* $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap i\mathbb{R} = \sigma_{\epsilon/\sqrt{2}}(\mathrm{L}) \cap i\mathbb{R}$.

Proof: Let $\lambda \in i\mathbb{R}$. Then by Theorem 3.3.5, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{L}) = \sqrt{2}\,\|\mathrm{L}(\lambda)x\|_2/\|(1,\lambda)\|_2$ for all x such that $\|x\|_2 = 1$. Hence taking minimum over $\|x\|_2 = 1$, we obtain the desired results.■

Theorem 3.5.6. *Let* $\mathbb{S}$ *be the space of* H*-palindromic matrix pencils and* $\mathrm{L} \in \mathbb{S}$. *Let* $\lambda \in \mathbb{T}$. *Then for the spectral and the Frobenius norms on* $\mathbb{C}^{n\times n}$, *we have* $\eta^{\mathbb{S}}(\lambda, \mathrm{L}) = \eta(\lambda, \mathrm{L})$ *and* $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap \mathbb{T} = \sigma_\epsilon(\mathrm{L}) \cap \mathbb{T}$.

Proof: Let L be given by $\mathrm{L}(\lambda) = A + \lambda A^H$. For $\lambda \in \mathbb{T}$, we have $\mathrm{L}(\lambda)^H = \overline{\lambda}\mathrm{L}(\lambda)$. This shows that $\mathrm{L}(\lambda)$ is a normal matrix. Let (μ, u) be an eigenpair of $\overline{\lambda}\mathrm{L}(\lambda)$ such that $|\mu| = \sigma_{\min}(\overline{\lambda}\mathrm{L}(\lambda)) = \sigma_{\min}(\mathrm{L}(\lambda))$. Define $\triangle A := -\frac{1}{2}\lambda\mu\, uu^H$ and consider the pencil $\triangle\mathrm{L}(z) = \triangle A + z(\triangle A)^H$. Noting the fact that $\overline{\lambda}\mathrm{L}(\lambda)u = \mu u$ and $\overline{\mu}u = (\overline{\lambda}\mathrm{L}(\lambda))^H u = \mu u$, we have $\mathrm{L}(\lambda)u + \triangle\mathrm{L}(\lambda)u = \lambda\mu\, u - \lambda\mu u = 0$. Further, we have $\|\triangle\mathrm{L}\| = |\mu|/\sqrt{2} = \sigma_{\min}(\mathrm{L}(\lambda))/\|(1,\lambda)\|_2 = \eta(\lambda, \mathrm{L})$. Hence the results follow. ■

For structured pencils, we have seen that $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap \Omega = \sigma_\epsilon(\mathrm{L}) \cap \Omega$ for appropriate $\Omega \subset \mathbb{C}$. We now show that this result plays an important role in solving certain distance problems associated with structured pencils. For illustration, we consider an H-even pencil $\mathrm{L}(z) = A + zB$. Then by Theorem 3.5.4, we have $\Omega = i\mathbb{R}$, that is, $\sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap i\mathbb{R} = \sigma_\epsilon(\mathrm{L}) \cap i\mathbb{R}$. The spectrum of L has Hamiltonian eigensymmetry, that is, the eigenvalues of L occur in $\lambda, -\overline{\lambda}$ pairs so that the eigenvalues are symmetric with respect to the imaginary axis $i\mathbb{R}$.

Question: *Suppose that* L *is* H*-even and is of size* $2n$. *Suppose also that* L *has* n *eigenvalues in the open left half complex plane and* n *eigenvalues in the open right half complex plane. What is the smallest value of* $\|\triangle\mathrm{L}\|$ *such that* $\triangle\mathrm{L}$ *is* H*-even and* $\mathrm{L} + \triangle\mathrm{L}$ *has a purely imaginary eigenvalue?*

Distance problems of this kind occur in many applications (see, for example, [30]). Let $\mathrm{d}(\mathrm{L})$ denote the smallest value of $\|\triangle\mathrm{L}\|$ such that $\mathrm{L} + \triangle\mathrm{L}$ has a purely imaginary eigenvalue. Then by Theorem 3.5.4, we have

$$\mathrm{d}(\mathrm{L}) = \inf_{t\in\mathbb{R}} \eta^{\mathbb{S}}(it, \mathrm{L}) = \min\{\epsilon : \sigma_\epsilon^{\mathbb{S}}(\mathrm{L}) \cap i\mathbb{R} \neq \emptyset\} = \min\{\epsilon : \sigma_\epsilon(\mathrm{L}) \cap i\mathbb{R} \neq \emptyset\} = \inf_{t\in\mathbb{R}} \eta(it, \mathrm{L}).$$

Hence $\mathrm{d}(\mathrm{L})$ can be read off from the unstructured pseudospectra of L. Note that $\eta(z, \mathrm{L}) = \sigma_{\min}(A + zB)/\sqrt{1+|z|^2}$. Thus if the infimum of $\eta(z, \mathrm{L})$ is attained at $\mu \in i\mathbb{R}$ then as in the proof of Theorem 3.5.4 we can construct an H-even pencil $\triangle\mathrm{L}$ such that μ is an eigenvalue of $\mathrm{L} + \triangle\mathrm{L}$ and that $\|\triangle\mathrm{L}\| = \eta(\mu, \mathrm{L}) = \mathrm{d}(\mathrm{L})$.

TH-789_04612301

Chapter 4

Structured backward errors and linearizations for structured matrix polynomials

In this chapter we derive explicit, computable expression of structured backward error of approximate eigenpair of structured regular matrix polynomials including symmetric, skew-symmetric, Hermitian, skew-Hermitian, even and odd. We also determine minimal structured perturbation such that the approximate eigenpair is the exact eigenepair of the perturbed polynomial. Using structured backward error expression for polynomial we obtain the same for widely varying structured linearizations of the polynomial. This yields suitable bound of its ratio with the unstructured backward error of approximate eigenpair of the polynomial. Thus we identify a "good" structured linearization which provides a negligible increment of structured backward error than the unstructured backward error of an approximate eigenpair of the polynomial. Finally we define the structured backward error of an approximate eigenvalue and this applied to establish a partial equality between unstructured and structured pseudospectra of the given structured polynomial.

4.1 Introduction

The polynomial eigenvalue problem is concerned with finding $(\lambda, x, y) \in \mathbb{C} \times \mathbb{C}^n \times \mathbb{C}^n$ which satisfies $\mathrm{P}(\lambda)x = 0$ and $y^H\mathrm{P}(\lambda) = 0$ where

$$\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j, \ A_j \in \mathbb{C}^{n\times n}, \text{ and } A_m \neq 0$$

is a matrix polynomial of degree m. In our study we assume that P is a regular, that is, $\det(\mathrm{P}(\lambda)) \neq 0$ for some scalar $\lambda \in \mathbb{C}$. The standard way to solve this problem is to convert P into an equivalent linear polynomial, called a linearization of P,

$$\mathrm{L}(\lambda) = \lambda X + Y, \ X, Y \in \mathbb{C}^{mn\times mn}$$

and employ a numerically backward stable algorithm to compute the eigenelements of L.

TH-789_04612301

It is well known that eigenvalues of a structured matrix polynomial inherit a spectral symmetry and that the spectral symmetry often has some physical significance [22, 68, 82, 83]. Therefore the crucial task for an algorithm to solving a structured polynomial eigenvalue problem is to preserve the spectral symmetries in the computed eigenvalues. There are a few structured preserving algorithms available in the literature, see [9, 10, 18, 44, 46, 71, 74, 82] (and the reference therein). To reveal the stability of such algorithms it is desirable to have the explicit expression of structured backward error of approximate eigenelements of a structured polynomial. We undertake a detailed backward perturbation analysis of structured matrix polynomials and their structured linearizations.

Example of structures that we consider in this chapter are T-symmetric, T-skew-symmetric, T-odd, T-even, H-Hermitian, H-skew-Hermitian, H-even and H-odd. In particular, our structures include polynomials whose coefficient matrices are Hamiltonians and skew-Hamiltonians. We denote the space of structured polynomials by $\mathbb{S}$ and we equip appropriate norm $|||\cdot|||$ on $\mathbb{S}$. Given a structured polynomial $\mathrm{P} \in \mathbb{S}$ and $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ with $x^H x = 1$, we determine the structured backward error $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ of (λ, x) as an approximate eigenpair of $\mathrm{P} \in \mathbb{S}$ and construct a polynomial $\triangle\mathrm{P} \in \mathbb{S}$ such that $|||\triangle\mathrm{P}||| = \eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ and $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$. Moreover, we show that $\triangle\mathrm{P}$ is unique for the Frobenius norm on $\mathbb{C}^{n\times n}$ but there are infinitely many such $\triangle\mathrm{P}$ for the spectral norm on $\mathbb{C}^{n\times n}$. Further, for the spectral norm, we show how to construct all such $\triangle\mathrm{P}$.

As said before, the first step towards solving a polynomial eigenvalue problem is linearization. For a structured matrix polynomial the job is to linearize the polynomial in such a way that the linearization reflects the structure of the polynomial and preserves the spectral symmetry. It is shown in [40, 68] that a structured matrix polynomial admits a plenty of structured linearizations. This poses a genuine problem of choosing one linearization over other. However for computational purposes, it is highly desirable to investigate how different structured linearizations affect the accuracy of computed eigenvalues.

In a view to analyze this we consider possible structured linearizations of a given $\mathrm{P} \in \mathbb{S}$. Then we derive structured backward error $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L})$ of approximate eigenpair $(\lambda, \Lambda_{m-1} \otimes x)$ of widely varying structured linearizations L, where $\Lambda_{m-1} := [\lambda^{m-1}, \ldots, \lambda, 1]^T$. We identify a potential *good* structured linearization L of P that minimizes $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L})/\eta(\lambda, x, \mathrm{P})$.

Moreover we consider structured pseudospectra of structured matrix polynomials. Pseudospectra of matrix polynomials have been studied extensively over the years (see, for example, [1, 2, 100] and the references therein). We define structured backward error of an approximate eigenvalue of a structured matrix polynomial and analyze structured pseudospectra. Then we establish a partial equality between structured and unstructured pseudospectra of structured polynomials.

The chapter is organized as follows. In section 4.2, we first define the structured polynomials and present the eigen-symmetry of these polynomials. In section 4.3, we obtain the expressions of structured backward error of an approximate eigenpair of structured polynomials. In section 4.4 we explain the structured linearizations of structure polynomials and determine the 'good' linearizations. The last section 4.5 we apply the expressions of structured backward errors to find the structured pseudospectra inclusions.

TH-789_04612301

4.2 Structured matrix polynomials

Recall that $\mathbb{P}_m(\mathbb{C}^{n\times n})$ denotes the space of matrix polynomials of degree m. Let $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j \in \mathbb{P}_m(\mathbb{C}^{n\times n})$. We say that (λ, x, y) is an eigentriple of P if λ is an eigenvalue of P and x and y, are the corresponding nonzero right and left eigenvectors respectively. We say $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ to be a structured polynomial if $\mathrm{P} \in \mathbb{S}$, where $\mathbb{S}$ is defined in Table 4.1. We define a map $\mathbb{P}_m(\mathbb{C}^{n\times n}) \to \mathbb{P}_m(\mathbb{C}^{n\times n})$, $\mathrm{P} \mapsto \mathrm{P}^*$ given by $\mathrm{P}^*(z) = \sum_{j=0}^{m} z^j A_j^*$, where $A^* = A^T$ or $A^* = A^H$. If $\mathrm{P} \in \mathbb{S}$ then P satisfies the condition given in the second column of Table 4.1.

$\mathbb{S}$	**Condition**	**eigen-symmetry**	**eigentriple**
T-symmetric	$\mathrm{P}^T(z) = \mathrm{P}(z), \forall z \in \mathbb{C}$	λ	$(\lambda, x, \overline{x})$
T-skew-symmetric	$\mathrm{P}^T(z) = -\mathrm{P}(z), \forall z \in \mathbb{C}$		
T-even	$\mathrm{P}^T(z) = \mathrm{P}(-z), \forall z \in \mathbb{C}$	$(\lambda, -\lambda)$	$(\lambda, x, \overline{y}), (-\lambda, y, \overline{x})$
T-odd	$\mathrm{P}^T(z) = -\mathrm{P}(-z), \forall z \in \mathbb{C}$		
H-Hermitian	$\mathrm{P}^H(z) = \mathrm{P}(z), \forall z \in \mathbb{C}$	$(\lambda, \overline{\lambda})$	$(\lambda, x, y), (\overline{\lambda}, y, x)$
H-skew-hermitian	$\mathrm{P}^H(z) = -\mathrm{P}(z), \forall z \in \mathbb{C}$		
H-even	$\mathrm{P}^H(z) = \mathrm{P}(-z), \forall z \in \mathbb{C}$	$(\lambda, -\overline{\lambda})$	$(\lambda, x, y), (-\overline{\lambda}, y, x)$
H-odd	$\mathrm{P}^H(z) = -\mathrm{P}(-z), \forall z \in \mathbb{C}$		

Table 4.1: Eigen-symmetry of structured polynomials.

Notice that H-Hermitian, H-skew-Hermitian polynomials have the same eigen-symmetry and $*$-even and $*$-odd polynomials, $* \in \{T, H\}$, have the same eigen-symmetry. It is easy to see that $\mathbb{S} \subset \mathbb{P}_m(\mathbb{C}^{n\times n})$ is a real/complex linear subspace of $\mathbb{P}_m(\mathbb{C}^{n\times n})$. Now we equip norms on $\mathbb{P}_m(\mathbb{C}^{n\times n})$ and make it normed linear space. For $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ we consider the polynomial norms

$$\|\mathrm{P}\|_F := \left(\sum_{j=0}^{m} \|A_j\|_F^2 \right)^{1/2} \quad \text{and} \quad \|\mathrm{P}\|_2 := \left(\sum_{j=0}^{m} \|A_j\|_2^2 \right)^{1/2}.$$

See [1, 2] for more on norms of matrix polynomials. We follow the convention that if $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is of the form $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$ then $\triangle\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is of the form $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Now we show that for any given $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $\mathrm{P} \in \mathbb{S}$ there always exists a polynomial $\triangle\mathrm{P} \in \mathbb{S}$ such that (λ, x) is a right eigenpair of $\mathrm{P} + \triangle\mathrm{P}$, that is, $(\mathrm{P}(\lambda) + \triangle\mathrm{P}(\lambda))x = 0$. For an $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$, we define the projection $P_x := I - xx^H$.

Theorem 4.2.1. *Let $\mathbb{S}$ be the space of structured matrix polynomials and $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. Assume $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ such that $x^H x = 1$. Set $r = -\mathrm{P}(\lambda)x$ and*

TH-789_04612301

$\Lambda_m := [1, \lambda, \ldots, \lambda^m]^T$. *Define*

$$\triangle A_j := \begin{cases} -\overline{x}x^T A_j x x^H + \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T + rx^H - 2(r^T x)\overline{x}x^H], & \text{if } A_j = A_j^T, \\ -\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T - rx^H], & \text{if } A_j = -A_j^T, \end{cases}$$

$$\triangle A_j := \begin{cases} -xx^H A_j x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\lambda^j x r^H P_x + \overline{\lambda^j} P_x r x^H], & \text{if } A_j = A_j^H, \\ -xx^H A_j x x^H - \frac{1}{\|\Lambda_m\|_2^2}[\lambda^j x r^H P_x - \overline{\lambda^j} P_x r x^H], & \text{if } A_j = -A_j^H, \end{cases}$$

and consider the polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$. *Then* $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$.

Proof: The proof is computational and is easy to check.■

4.3 Structured backward error

Recall that the spectrum of a regular polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is denoted by $\sigma(\mathrm{P})$, (chapter 1). In this chapter, we consider only finite eigenvalues of matrix polynomials. By convention, if $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ then x is assumed to be nonzero, that is, $x \neq 0$. Treating (λ, x) as an approximate eigenpair of the polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, we define the backward error of (λ, x) by

$$\eta_F(\lambda, x, \mathrm{P}) := \inf_{\triangle \mathrm{P}\in\mathbb{P}_m(\mathbb{C}^{n\times n})} \{\|\triangle\mathrm{P}\|_F : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}$$
$$\eta_2(\lambda, x, \mathrm{P}) := \inf_{\triangle \mathrm{P}\in\mathbb{P}_m(\mathbb{C}^{n\times n})} \{\|\triangle\mathrm{P}\|_2 : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}.$$

Further, for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r := -\mathrm{P}(\lambda)x$, we have

$$\eta_F(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|x\|_2\|\Lambda_m\|_2} = \eta_2(\lambda, x, \mathrm{P}). \tag{4.1}$$

Indeed, defining $\triangle A_j := \dfrac{\overline{\lambda^j} r x^H}{x^H x\|\Lambda_m\|_2^2}$, $j = 0 : m$, and considering the polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$, we have $\|\triangle\mathrm{P}\|_F = \|r\|_2/\|x\|_2\|\Lambda_m\|_2 = \|\triangle\mathrm{P}\|_2$ and $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$. Henceforth, we denote the unstructured backward error with respect to both Frobenius and spectral norms by $\eta(\lambda, x, \mathrm{P})$.

Next assume that $\mathrm{P} \in \mathbb{S}$. Then treating $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^{n\times n}$ as an approximate eigenpair of P, we define the structured backward error of (λ, x) by

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) := \inf_{\triangle\mathrm{P}\in\mathbb{S}} \{\|\triangle\mathrm{P}\|_F : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}$$
$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) := \inf_{\triangle\mathrm{P}\in\mathbb{S}} \{\|\triangle\mathrm{P}\|_2 : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}.$$

Henceforth, we denote the structured backward error with respect to both Frobenius and spectral norms by $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Obviously, we have $\eta(\lambda, x, \mathrm{P}) \leq \eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ and by Theorem 4.2.1, $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) < \infty$.

Now we derive the structured backward error of an approximate eigenpair (λ, x) of structured matrix polynomials. Recall that for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, our standing assumption is that

TH-789_04612301

$x^H x = 1$. First, consider T-symmetric matrix polynomials. Recall that $\Lambda_m := [1, \lambda, \ldots, \lambda^m]^T$.

Theorem 4.3.1. *Let $\mathbb{S}$ be the space of T-symmetric matrix polynomials and $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r := -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\sqrt{2\|r\|_2^2 - |x^T r|^2}}{\|\Lambda_m\|_2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), \;\; \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P}).$$

Now define $\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x} r^T + r x^H - (r^T x)\overline{x} x^H]$, $j = 0 : m$ and consider the polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Then $\triangle \mathrm{P}$ is a unique polynomial such that $\triangle \mathrm{P} \in \mathbb{S}$, $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Further, define

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x} r^T + r x^H - (r^T x)\overline{x} x^H] - \frac{\overline{\lambda^j}\; \overline{x^T r}\; P_x^T r r^T P_x}{\|\Lambda_m\|_2^2\;(\|r\|_2^2 - |x^T r|^2)}$$

and consider the polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Then $\triangle \mathrm{P} \in \mathbb{S}$, $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: By Theorem 4.2.1 there exists $\triangle \mathrm{P} \in \mathbb{S}$ such that $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$. Then we have $r = \triangle \mathrm{P}(\lambda)x$. Choose $Q_1 \in \mathbb{C}^{n \times (n-1)}$ such that the matrix $Q = [x \;\; Q_1]$ is unitary. Let $\widetilde{\triangle A_j} := Q^T \triangle A_j Q = \begin{pmatrix} a_{jj} & a_j^T \\ a_j & X_j \end{pmatrix}$, where $X_j = X_j^T$ is of size $n-1$. Since $\overline{Q} Q^T = I$, we have

$$\overline{Q}(\triangle \mathrm{P}(\lambda))Q^H x = r \Rightarrow (\triangle \mathrm{P}(\lambda))Q^H x = Q^T r = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}$$

As $Q^H x = e_1$, the first column of the identity matrix, we have $\begin{pmatrix} \sum_{j=0}^{m} \lambda^j a_{jj} \\ \sum_{j=0}^{m} \lambda^j a_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}$ whose minimum solutions are $a_j = \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2}, a_{jj} = \frac{\overline{\lambda^j} x^T r}{\|\Lambda_m\|_2^2}$, $j = 0 : m$. Hence we have

$$\widetilde{\triangle A_j} = \begin{pmatrix} \frac{\overline{\lambda^j} x^T r}{\|\Lambda_m\|_2^2} & (\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2})^T \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix}. \tag{4.2}$$

This shows that the Frobenius norm of $\widetilde{\triangle A_j}$'s are minimized when $X_j = 0$. Hence we have $\|\triangle A_j\|_F^2 = \|\widetilde{\triangle A_j}\|_F^2 = |a_{jj}|^2 + 2\|a_j\|_2^2$. Since $\overline{Q}_1 Q_1^T = I - \overline{x} x^T$, we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{|x^T r|^2}{\|\Lambda_m\|_2^2} + \frac{2\|(I - \overline{x} x^T) r\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\sqrt{2\|r\|_2^2 - |x^T r|^2}}{\|\Lambda_m\|_2}.$$

TH-789_04612301

Simplifying the expressions of $\triangle A_j$ we have

$$\begin{aligned}\triangle A_j &= [\overline{x}\ \overline{Q}_1]\begin{pmatrix} \frac{\overline{\lambda^j} x^T r}{\|\Lambda_m\|_2^2} & (\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2})^T \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & 0 \end{pmatrix}\begin{pmatrix} x^H \\ Q_1^H \end{pmatrix} \\ &= \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}(x^T r)\overline{x}x^H + \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}\overline{x}r^T Q_1 Q_1^H + \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}\overline{Q}_1 Q_1^T r x^H \\ &= \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}x^T r x^H + \overline{x}r^T(I - xx^H) + (I - \overline{x}x^T)rx^H] \\ &= \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T + rx^H - (r^T x)\overline{x}x^H]\end{aligned}$$

from which we obtain the desired polynomial $\triangle \mathrm{P}$.

From (4.2) consider $\mu_{\triangle A_j} = \frac{|\lambda^j|\ \|r\|_2}{\|\Lambda_m\|_2^2}$. Then by DKW Theorem 1.2.5 we have,

$$\triangle X_j = -\frac{\overline{\lambda^j}\ \overline{x^T r}\ Q_1^T r (Q_1^T r)^T}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)},\ j = 0 : m,$$

which gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|\Lambda_m\|_2} = \eta(\lambda, x, \mathrm{P}).$$

Simplifying the expression of $\triangle A_j$ we obtain

$$\triangle A_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T + rx^H - (r^T x)\overline{x}x^H] - \frac{\overline{\lambda^j}\ \overline{x^T r}\ P_x^T r r^T P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}.$$

This completes the proof. ■

Remark 4.3.2. *If $|x^T r| = \|r\|_2$, then $\|Q_1^T r\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

Observe that if Y is symmetric and $Yx = 0$ then $Y = P_x^T Z P_x$ for some symmetric matrix Z. Consequently, we have $\overline{Q_j} X_j Q_j^H = P_x^T Z_j P_x, j = 0 : m$, for some symmetric matrices Z_j. Hence from the proof of Theorem 4.3.1 we have following.

Corollary 4.3.3. *Let $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$ be a T-symmetric polynomial and let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Set $r = -\mathrm{P}(\lambda)x$. Then $\mathrm{P}(\lambda)x + \mathrm{Q}(\lambda)x = 0$ if and only if $\mathrm{Q}(z) = \triangle\mathrm{P}(z) + P_x^T \mathrm{R}(z) P_x$ for some T-symmetric polynomial $\mathrm{R} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, where $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is the T-symmetric polynomial, given by*

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T + rx^H - (r^T x)\overline{x}x^H]$$

Next we consider T-skew-symmetric polynomials.

Theorem 4.3.4. *Let $\mathbb{S}$ be the space of T-skew-symmetric matrix polynomials and $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r = -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\ \eta(\lambda, x, \mathrm{P}),\ \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P}).$$

TH-789_04612301

Further, for the T-skew-symmetric polynomial $\triangle\mathrm{P}$ *given in Theorem 4.2.1 we have* $(\mathrm{P}(\lambda)+\triangle\mathrm{P}(\lambda))x = 0$, $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$ *and* $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: The arguments proceed on the lines as those given in the proof of Theorem 4.3.1. The only difference is the fact that, in this case, $\triangle A_j$ is skew-symmetric for all $j = 0 : m$. Thus we have

$$\widetilde{\triangle A_j} := Q^T \triangle A_j Q = \begin{pmatrix} 0 & a_j^T \\ -a_j & X_j \end{pmatrix}, \quad Q^T r = \begin{pmatrix} 0 \\ Q_1^T r \end{pmatrix}, X_j^T = -X_j.$$

Consequently, we have

$$\begin{pmatrix} 0 \\ -\sum_{j=0}^m \lambda^j a_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix} \Rightarrow x^T r = 0,\ a_j = -\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2}.$$

Hence we have

$$\triangle A_j = \overline{Q} \begin{pmatrix} 0 & -(\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2})^T \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H. \tag{4.3}$$

Setting $X_j = 0$, we obtain the polynomial $\triangle\mathrm{P}$ such that $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\|r\|_2/\|\Lambda_m\|_2$. Next, since $\overline{Q}_1 Q_1^T = \mathrm{I} - \overline{x} x^T$, simplifying the expressions we have $\triangle A_i = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[r x^H - (r x^H)^T]$, from which the T-skew-symmetric polynomial $\triangle\mathrm{P}$ given in Theorem 4.2.1 follows. This completes the proof for Frobenius norm.

Next from (4.3) we have $\mu_{\triangle A_j} = \frac{|\lambda^j|\ \|r\|_2}{\|\Lambda_m\|_2^2}$. Hence by DKW Theorem 1.2.5, we have, $X_j = 0$. Thus we obtain $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P})$. The desired result follows by simplifying the expression of $\triangle A_j$. This completes the proof.■

Using the fact that if Y is skew-symmetric and $Yx = 0$ then $Y = P_x^T Z P_x$ for some skew-symmetric matrix Z, we obtain an analogue of Corollary 4.3.3 for T-skew-symmetric polynomials.

Corollary 4.3.5. *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ *be a T-skew-symmetric polynomial and let* $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$. *Set* $r = -\mathrm{P}(\lambda)x$. *Then* $\mathrm{P}(\lambda)x + \mathrm{Q}(\lambda)x = 0$ *if and only if* $\mathrm{Q}(z) = \triangle\mathrm{P}(z) + P_x^T \mathrm{R}(z) P_x$ *for some T-skew-symmetric polynomial* $\mathrm{R} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, *where* $\triangle\mathrm{P}$ *is the T-skew-symmetric polynomial given in Theorem 4.3.4.*

To describe the structured backward errors for T-even and T-odd polynomials in a convenient manner, we define the even index projection $\Pi_e : \mathbb{C}^{m+1} \to \mathbb{C}^{m+1}$ by

$$\Pi_e([x_0, x_1, x_1, \ldots, x_{m-1}, x_m]^T) := \begin{cases} [x_0, 0, x_2, 0, \ldots, x_{m-2}, 0, x_m]^T, & \text{if } m \text{ is even,} \\ [x_0, 0, x_2, 0, \ldots, 0, x_{m-1}, 0]^T, & \text{if } m \text{ is odd.} \end{cases}$$

Note that "0" is considered as even number. Then $I - \Pi_e$ is the odd index projection.

Theorem 4.3.6. *Let* $\mathbb{S}$ *be the space of T-even matrix polynomials and* $\mathrm{P} \in \mathbb{S}$ *be given by* $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$. *Then for* $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$, *setting* $r = -\mathrm{P}(\lambda)x$, *we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{|x^T r|^2}{\|\Pi_e(\Lambda_m)\|_2^2} + 2\frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}},\ \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{|x^T r|^2}{\|\Pi_e(\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}.$$

TH-789_04612301

In particular, if m is odd and $|\lambda| = 1$, then $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\ \eta(\lambda, x, \mathrm{P})$, $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P})$.

Let

$$E_j := \begin{cases} \dfrac{\overline{\lambda^j}}{\|\Pi_e(\Lambda_m)\|_2^2}(x^T r)\overline{x}x^H + \dfrac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x}r^T P_x + P_x^T r x^H], & \text{if } j \text{ is even} \\ \dfrac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[P_x^T r x^H - \overline{x}r^T P_x], & \text{if } j \text{ is odd .} \end{cases}$$

Setting $\triangle A_j = E_j$ we obtain a unique T-even polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Further, for $j = 0 : m$ defining

$$\triangle A_j := \begin{cases} E_j - \dfrac{\overline{\lambda^j}\ \overline{x^T r}\ P_x^T r r^T P_x}{\|\Pi_e(\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } j \text{ is even} \\ E_j, & \text{if } j \text{ is odd} \end{cases}$$

we obtain a T-even polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: Suppose $\mathbb{S}$ is a space of T-even polynomials and $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j \in \mathbb{S}$. Note that A_j is symmetric when j is even (including "0") and skew-symmetric when j is odd. The proof follows from similar arguments as those employed for T-symmetric and T-skew-symmetric polynomials. Indeed, considering a unitary matrix $Q := [x,\ Q_1]$, we have $\triangle A_j = \overline{Q}\begin{pmatrix} a_{jj} & a_j^T \\ a_j & X_j \end{pmatrix} Q^H$, $X_j^T = X_j$, if j is even, and $\triangle A_j = \overline{Q}\begin{pmatrix} 0 & b_j^T \\ -b_j & Y_j \end{pmatrix} Q^H$, $Y_j^T = -Y_j$, if j is odd.

Consequently we have

$$\begin{pmatrix} \sum_j \lambda^j a_{jj} \\ \sum_{j\text{-even}} \lambda^j a_j - \sum_{j\text{-odd}} \lambda^j b_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

Hence the smallest norm solutions are $a_{jj} = \frac{\overline{\lambda^j}}{\|\Pi_e(\Lambda_m)\|_2^2}\, x^T r$, $a_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r$, $b_j = -\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r$. Therefore we have

$$\triangle A_j = \begin{cases} \overline{Q}\begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Pi_e(\Lambda_m)\|_2^2} x^T r & (\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2})^T \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, & \text{if } j \text{ is even} \\ \overline{Q}\begin{pmatrix} 0 & -(\frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2})^T \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & Y_j \end{pmatrix} Q^H, & \text{if } j \text{ is odd.} \end{cases} \tag{4.4}$$

Setting $X_j = 0 = Y_j$ and using the fact that $\overline{Q}_1 Q_1^T = \mathrm{I} - \overline{x}x^T$, we obtain a unique polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ such that

$$\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{|x^T r|^2}{\|\Pi_e(\Lambda_m)\|_2^2} + 2\frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}.$$

TH-789_04612301

If m is odd and $|\lambda| = 1$ then notice that $\|\Pi_e(\Lambda_m)\|_2^2 = \frac{1}{2}\|\Lambda_m\|_2^2$. Hence we obtain $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\,\eta(\lambda, x, \mathrm{P})$. Now simplifying expressions for $\triangle A_j$ we obtain $\triangle A_j = E_j$. Therefore, we obtain a T-even polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

For the spectral norm, we consider $\mu_{\triangle A_j} = \sqrt{\frac{|\lambda^j|^2\ |x^Tr|^2}{\|\Pi_e(\Lambda_m)\|_2^4} + \frac{|\lambda^j|^2\ (\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}$, if j is even and $\mu_{\triangle A_j} = \sqrt{\frac{|\lambda^j|^2\ (\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}$, if j is odd.

Then by DKW Theorem 1.2.5 and (4.4) we have,

$$X_j = -\frac{\overline{\lambda^j}\ \overline{x^Tr}\ Q_1^T r (Q_1^T r)^T}{\|\Pi_e(\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^Tr|^2)} \text{ and } Y_j = 0$$

which gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{|x^Tr|^2}{\|\Pi_e(\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^Tr|^2}{\|\Lambda_m\|_2^2}}.$$

Now simplifying expressions for $\triangle A_j$ we obtain the desired result.■

The above proof shows that setting $\triangle A_j := E_j + P_x^T Z_j P_x$, where $Z_j^T = Z_j$ when j is even and $Z_j^T = -Z_j$ when j is odd, we obtain a T-even polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$.

Remark 4.3.7. *If $|x^Tr| = \|r\|_2$, then $\|Q_1^T r\|_2 = 0$. In such a case, considering $X_j = 0 = Y_j$ we obtain the desired results.*

Next we consider backward error of T-odd polynomials.

Theorem 4.3.8. *Let $\mathbb{S}$ be the space of T-odd matrix polynomials and $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r = -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{\frac{|x^Tr|^2}{\|(I-\Pi_e)(\Lambda_m)\|_2^2} + 2\frac{\|r\|_2^2 - |x^Tr|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq 0 \\ \sqrt{2}\,\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = 0 \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{\frac{|x^Tr|^2}{\|(I-\Pi_e)(\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^Tr|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq 0 \\ \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = 0 \end{cases}$$

In particular, if m is odd and $|\lambda| = 1$, then $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\ \eta(\lambda, x, \mathrm{P})$, $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P})$.

Let

$$F_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[P_x^T r x^H - \overline{x} r^T P_x], & \text{if } j \text{ is even} \\ \frac{\overline{\lambda^j}\overline{x}x^T r x^H}{\|(I-\Pi_e)(\Lambda_m)\|_2^2} + \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[\overline{x} r^T P_x + P_x^T r x^H], & \text{if } j \text{ is odd.} \end{cases}$$

Defining $\triangle A_j := F_j$, we obtain a unique T-odd polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Further for the spectral norm, define $\triangle A_j := F_j$, if j is even and

$$\triangle A_j := F_j - \frac{\overline{\lambda^j}\ \overline{x^Tr} P_x^T r r^T P_x}{\|(I-\Pi_e)\Lambda_m\|_2^2(\|r\|_2^2 - |x^Tr|^2)},$$

TH-789_04612301

if j is odd. Then we obtain a T-odd polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: By interchanging the role of A_j for even and odd j the desired result follows from the proof of Theorem 4.3.6. ■

The above Theorem shows that setting $\triangle A_j = E_j + P_x^T Z_j P_x$, where $Z_j^T = -Z_j$ when j is even, and $Z_j^T = Z_j$ when j is odd, we obtain a T-odd polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ such that $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$.

To make the presentation simple we introduce the following operators.

$\mathsf{Re} : \mathbb{C}^{m+1} \to \mathbb{R}^{m+1}$ is defined by $\mathsf{Re}([x_0, x_1, \ldots, x_m]^T) \mapsto [\mathsf{re}(x_0), \mathsf{re}(x_1), \ldots, \mathsf{re}(x_m)]^T$

$\mathsf{Im} : \mathbb{C}^{m+1} \to \mathbb{R}^{m+1}$ is defined by $\mathsf{Im}([x_0, x_1, \ldots, x_m]^T) \mapsto [\mathsf{im}(x_0), \mathsf{im}(x_1), \ldots, \mathsf{im}(x_m)]^T$.

Then for $x \in \mathbb{C}^{m+1}$ we have $x = \mathsf{Re}(x) + i\,\mathsf{Im}(x)$, where $i = \sqrt{-1}$ and $\mathsf{re}(z), \mathsf{im}(z)$ denote the real and imaginary parts of a complex number z, respectively.

Theorem 4.3.9. *Let $\mathbb{S}$ be the space of H-Hermitian or H-skew-Hermitian matrix polynomials and $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r = -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|\Lambda_m\|_2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in \mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{2(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus \mathbb{R}. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in \mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus \mathbb{R}. \end{cases}$$

where $\widehat{r} = \begin{bmatrix} \mathsf{Re}(\Lambda_m)^T \\ \mathsf{Im}(\Lambda_m)^T \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{bmatrix}$ for H-Hermitian and $\widehat{r} = \begin{bmatrix} -\mathsf{Im}(\Lambda_m)^T \\ \mathsf{Re}(\Lambda_m)^T \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{bmatrix}$ for H-skew-Hermitian polynomial.

Next, let $E = xr^H + rx^H - (r^H x)xx^H$ and $F = rx^H - xr^H + (r^H x)xx^H$.

When $\lambda \in \mathbb{R}$, define

$$\triangle A_j := \begin{cases} \frac{\lambda^j}{\|\Lambda_m\|_2^2} E, & \text{if } A_j = A_j^H \\ \frac{\lambda^j}{\|\Lambda_m\|_2^2} F, & \text{if } A_j = -A_j^H \end{cases}$$

and for $\lambda \in \mathbb{C} \setminus \mathbb{R}$, define

$$\triangle A_j := \begin{cases} e_j^T \widehat{r} x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H + \lambda^j x r^H P_x], & \text{if } A_j = A_j^H \\ i e_j^T \widehat{r} x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H - \lambda^j x r^H P_x], & \text{if } A_j = -A_j^H. \end{cases}$$

Now Consider the polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Then $\triangle \mathrm{P} \in \mathbb{S}$ is unique such that $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

TH-789_04612301

Further, for $\lambda \in \mathbb{R}$, define

$$\triangle A_j := \begin{cases} \dfrac{\lambda^j}{\|\Lambda_m\|_2^2} E - \dfrac{\lambda^j \; x^H r P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } A_j = A_j^H \\ \dfrac{\lambda^j}{\|\Lambda_m\|_2^2} F + \dfrac{\lambda^j r^H x P_x r r^H P_x}{\|\Lambda_m\|_2^2 (\|r\|_2^2 - |x^H r|^2)}, & \text{if } A_j = -A_j^H. \end{cases}$$

and for $\lambda \in \mathbb{C} \setminus \mathbb{R}$, define

$$\triangle A_j := \begin{cases} e_j^T \widehat{r} x x^H + \dfrac{1}{\|\Lambda_m\|_2^2} [\overline{\lambda^j} P_x r x^H + \lambda^j x r^H P_x] - \dfrac{e_j^T \widehat{r} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } A_j = A_j^H \\ i e_j^T \widehat{r} x x^H + \dfrac{1}{\|\Lambda_m\|_2^2} [\overline{\lambda^j} P_x r x^H - \lambda^j x r^H P_x] + \dfrac{\overline{i e_j^T \widehat{r}} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } A_j = -A_j^H. \end{cases}$$

Consider the polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$. Then $\triangle \mathrm{P} \in \mathbb{S}$, $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda) x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: Suppose $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is an H-Hermitian matrix polynomial. By Theorem 4.2.1 there exists an H-Hermitian matrix polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ such that $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. Now choosing a unitary matrix $Q := [x, \; Q_1]$, we have

$$\widetilde{\triangle A_j} := Q^H \triangle A_j Q = \begin{pmatrix} a_{jj} & a_j^H \\ a_j & X_j \end{pmatrix}, \quad Q^H r = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}.$$

Now $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0 \Rightarrow \begin{pmatrix} \sum_{j=0}^m \lambda^j a_{jj} \\ \sum_{j=0}^m \lambda^j a_j \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. The minimum norm solution of $\sum_{j=0}^m \lambda^j a_j = Q_1^H r$ is given by $a_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r$.

If $\lambda \in \mathbb{R}$ then minimum norm solution of $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$ is $a_{jj} = \frac{\lambda^j}{\|\Lambda_m\|_2^2} x^H r \in \mathbb{R}$. Hence for $\lambda \in \mathbb{R}$ we have

$$\triangle A_j = Q \begin{pmatrix} \frac{\lambda^j}{\|\Lambda_m\|_2^2} x^H r & (\frac{\lambda^j}{\|\Lambda_m\|_2^2} Q_1^H r)^H \\ \frac{\lambda^j}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H, \; j = 0 : m. \tag{4.5}$$

Setting $X_j = 0$ we obtain $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\sqrt{2\|r\|_2^2 - |r^H x|^2}}{\|\Lambda_m\|_2}$. Simplifying the expression of $\triangle A_j$ we obtain the desired result.

If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$ gives

$$\begin{pmatrix} \sum_{j=0}^m \mathsf{re}(\lambda^j) a_{jj} \\ \sum_{j=0}^m \mathsf{im}(\lambda^j) a_{jj} \end{pmatrix} = \begin{pmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{pmatrix} \Rightarrow \begin{pmatrix} a_{00} \\ \vdots \\ a_{mm} \end{pmatrix} = M^{\dagger} \begin{pmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{pmatrix} =: \widehat{r}(\text{say})$$

where $M = \begin{pmatrix} \mathsf{Re}(\Lambda_m)^T \\ \mathsf{Im}(\Lambda_m)^T \end{pmatrix}$. Therefore $a_{jj} = e_j^T \widehat{r}$, where e_j is the j-th column of identity

TH-789_04612301

matrix. Setting $X_j = 0$ we obtain

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\|\widehat{r}\|_2^2 + 2\frac{\|r\|_2^2 - |r^H x|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expressions of $\triangle A_i$ we obtain the desired result.

Next consider the spectral norm. Let $\lambda \in \mathbb{R}$. For $\mu_{\triangle A_j} = \frac{|\lambda^j| \, \|r\|_2}{\|\Lambda_m\|_2^2}$, $j = 0 : m$, by DKW Theorem 1.2.5 and (4.5) we obtain

$$X_j = -\frac{\lambda^j \; x^H r (Q_1^H r)(Q_1^H r)^H}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^H r|^2)}.$$

This gives, $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|\Lambda_m\|_2} = \eta(\lambda, x, \mathrm{P})$. Simplifying the expression of $\triangle A_i$ we obtain the desired result.

Now if $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then for $\mu_{\triangle A_j} = \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^j|^2 \; (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}$ applying the DKW Theorem 1.2.5 to $\triangle A_j$ we have

$$X_j = -\frac{e_j^T \widehat{r} \; (Q_1^H r)(Q_1^H r)^H}{\|r\|_2^2 - |x^H r|^2}, \; j = 0 : m.$$

This gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expression of $\triangle A_j$, $j = 0 : m$ we obtain

$$\triangle A_j = e_j^T \widehat{r} x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H + \lambda^j x r^H P_x] - \frac{e_j^T \widehat{r} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}.$$

The proof is similar when P is H-skew-Hermitian polynomial.■

Remark 4.3.10. *If $|x^H r| = \|r\|_2$, then $\|Q_1^H r\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

Theorem 4.3.11. *Let $\mathbb{S}$ be the space of H-even matrix polynomials. Let $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r = -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|\Lambda_m\|_2} \le \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{2(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R}. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R} \end{cases}$$

where $\widehat{r} = \begin{bmatrix} \Pi_e \, \mathsf{Re}(\Lambda_m)^T - (I - \Pi_e)\mathsf{Im}(\Lambda_m)^T \\ \Pi_e \, \mathsf{Im}(\Lambda_m)^T + (I - \Pi_e)\mathsf{Re}(\Lambda_m)^T \end{bmatrix}^\dagger \begin{bmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{bmatrix}.$

Set

$$E_j := \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H + \lambda^j x r^H P_x], \quad F_j := \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H - \lambda^j x r^H P_x]. \tag{4.6}$$

TH-789_04612301

When $\lambda \in i\,\mathbb{R}$, define

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[xr^H + rx^H - (r^H x)xx^H].$$

For $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, define

$$\triangle A_j := \begin{cases} e_j^T \widehat{r} x x^H + E_j, & \text{if } j \text{ is even} \\ i e_j^T \widehat{r} x x^H + F_j, & \text{if } j \text{ is odd.} \end{cases}$$

Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$. is a unique polynomial such that $\triangle \mathrm{P} \in \mathbb{S}$, $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Further, when $\lambda \in i\,\mathbb{R}$ define

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[xr^H + rx^H - (r^H x)xx^H] + \frac{(-1)^{j+1}\lambda^j \; x^H r P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^H r|^2)}$$

and for $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, define

$$\triangle A_j := \begin{cases} e_j^T \widehat{r} x x^H + E_j + \dfrac{(-1)^{j+1} e_j^T \widehat{r} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } j \text{ is even} \\ i e_j^T \widehat{r} x x^H + F_j + \dfrac{(-1)^{j+1} i e_j^T \widehat{r} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } j \text{ is odd.} \end{cases}$$

Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is such that $\triangle \mathrm{P} \in \mathbb{S}$, $\mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: By Theorem 4.2.1, there exists a H-even matrix polynomial $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ such that $\triangle \mathrm{P}(\lambda) = r$. Now choosing a unitary matrix $Q := [x, \; Q_1]$, and noting that $\triangle A_j = \triangle A_j^H$ when j is even, and $\triangle A_j = -\triangle A_j^H$ when j is odd, we have $\triangle A_j = Q \begin{pmatrix} a_{jj} & a_j^H \\ a_j & X_j \end{pmatrix} Q^H$, $X_j^H = X_j$ if j is even, and $\triangle A_j = Q \begin{pmatrix} i a_{jj} & a_j^H \\ -a_j & Y_j \end{pmatrix} Q^H$, $Y_j^H = -Y_j$ if j is odd. Notice that a_{jj} is real for all j.

Then $\triangle \mathrm{P}(\lambda)x = r$ gives $\begin{pmatrix} \sum_{j\text{-even}} \lambda^j a_{jj} + i \sum_{j\text{-odd}} \lambda^j a_{jj} \\ \sum_{j\text{-even}} \lambda^j a_j - \sum_{j\text{-odd}} \lambda^j a_j \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}$. The minimum norm solution of $\sum_{j\text{-even}} \lambda^j a_j - \sum_{j\text{-odd}} \lambda^j a_j = Q_1^H r$ is given by $a_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r$, if j is even and $a_j = -\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r$, if j is odd.

If $\lambda \in i\;\mathbb{R}$, then the minimum norm solution for a_{jj} are given by $a_{jj} = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r$ when j is even, and $a_{jj} = -\frac{i\,\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r$ when j is odd. Then $a_{jj} \in \mathbb{R}$, when j is even and $i a_{jj} \in i\,\mathbb{R}$ if j is odd. Hence if $\lambda \in i\,\mathbb{R}$, then we have

$$\triangle A_j = Q \begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r & (\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H$$

TH-789_04612301

when j is even, and

$$\triangle A_j = Q \begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r & -(\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & Y_j \end{pmatrix} Q^H$$

when j is odd. Setting $X_j = 0 = Y_j$, we obtain $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\sqrt{2\|r\|_2^2 - |r^H x|^2}}{\|\Lambda_m\|_2}$. Now simplifying the expressions for $\triangle A_j$ we obtain the desired result.

Next if $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, then $\sum_{j\text{-even}} \lambda^j a_{jj} + i \sum_{j\text{-odd}} \lambda^j a_{jj} = x^H r$ gives

$$\begin{pmatrix} \sum_{j\text{-even}} \mathsf{re}(\lambda^j) a_{jj} - \sum_{j\text{-odd}} \mathsf{im}(\lambda^j) a_{jj} \\ \sum_{j\text{-even}} \mathsf{im}(\lambda^j) a_{jj} + \sum_{j\text{-odd}} \mathsf{re}(\lambda^j) a_{jj} \end{pmatrix} = \begin{pmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{pmatrix}.$$

Hence we have

$$\begin{pmatrix} a_{00} \\ a_{11} \\ \vdots \\ a_{mm} \end{pmatrix} = K^\dagger \begin{pmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{pmatrix} =: \widehat{r}(say) \Rightarrow a_{jj} = e_j^T \widehat{r}$$

where $K = \begin{pmatrix} \Pi_e \, \mathsf{Re}(\Lambda_m)^T - (I - \Pi_e)\mathsf{Im}(\Lambda_m)^T \\ \Pi_e \, \mathsf{Im}(\Lambda_m)^T + (I - \Pi_e)\mathsf{Re}(\Lambda_m)^T \end{pmatrix}$. Consequently we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\|\widehat{r}\|_2^2 + 2\frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expression of $\triangle A_j,\ j = 0 : m$ we obtain the desired result.

Now we consider spectral norm. For $\lambda \in i\,\mathbb{R}$, consider $\mu_{\triangle A_j} = \frac{|\lambda^j| \ \|r\|_2}{\|\Lambda_m\|_2^2}$ if j is even, and $\mu_{\triangle A_j} = \frac{|\lambda^j| \ \|r\|_2}{\|\Lambda_m\|_2^2}$ when j is odd. Then by DKW Theorem 1.2.5 applied to $\triangle A_j$, gives

$$X_j = -\frac{\lambda^j \ x^H r (Q_1^H r)(Q_1^H r)^H}{\|\Lambda_m\|_2^2 \ (\|r\|_2^2 - |x^H r|^2)}, \quad Y_j = \frac{\lambda^j \ x^H r (Q_1^H r)(Q_1^H r)^H}{\|\Lambda_m\|_2^2 \ (\|r\|_2^2 - |x^H r|^2)}.$$

This gives $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|\Lambda_m\|_2}$. Simplifying the expressions for $\triangle A_j$ we obtain the desired result.

For $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, consider $\mu_{\triangle A_j} = \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^j|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}$, for $j = 0 : m$. Then by DKW Theorem 1.2.5 applied to $\triangle A_j$ gives

$$X_j = -\frac{e_j^T \widehat{r} \ (Q_1^H r)(Q_1^H r)^H}{\|r\|_2^2 - |x^H r|^2}, \quad Y_j = \frac{\overline{i e_j^T \widehat{r}} \ (Q_1^H r)(Q_1^H r)^H}{\|r\|_2^2 - |x^H r|^2}.$$

This gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expression of $\triangle A_j,\ j = 0 : m$ we obtain the desired result.■

TH-789_04612301

Theorem 4.3.12. *Let $\mathbb{S}$ be the space of H-odd matrix polynomials. Let $\mathrm{P} \in \mathbb{S}$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, setting $r = -\mathrm{P}(\lambda)x$, we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{\sqrt{2\|r\|_2^2 - |x^H r|^2}}{\|\Lambda_m\|_2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{2(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R}. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R} \end{cases}$$

where $\widehat{r} = \begin{bmatrix} -\Pi_e \,\mathsf{Im}(\Lambda_m)^T + (I - \Pi_e)\mathsf{Re}(\Lambda_m)^T \\ \Pi_e \,\mathsf{Re}(\Lambda_m)^T + (I - \Pi_e)\mathsf{Im}(\Lambda_m)^T \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(x^H r) \\ \mathsf{im}(x^H r) \end{bmatrix}.$

When $\lambda \in i\,\mathbb{R}$ define

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[xr^H - rx^H + (r^H x)xx^H].$$

For $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, define

$$\triangle A_j := \begin{cases} ie_j^T \widehat{r} x x^H + F_j, & \text{if } j \text{ is even} \\ e_j^T \widehat{r} x x^H + E_j, & \text{if } j \text{ is odd} \end{cases}$$

where E_j and F_j are given in (4.6). Then $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ is a unique polynomial such that $\triangle\mathrm{P} \in \mathbb{S}$, $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Further, when $\lambda \in i\,\mathbb{R}$ define

$$\triangle A_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}[xr^H - rx^H + (r^H x)xx^H] + \frac{(-1)^j \lambda^j \; x^H r P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^H r|^2)}.$$

When $\lambda \in \mathbb{C} \setminus i\,\mathbb{R}$, define

$$\triangle A_j := \begin{cases} ie_j^T \widehat{r} x x^H + F_j + \dfrac{(-1)^j \overline{ie_j^T \widehat{r}} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } j \text{ is even} \\[2ex] e_j^T \widehat{r} x x^H + E_j + \dfrac{(-1)^j e_j^T \widehat{r} \; P_x r r^H P_x}{\|r\|_2^2 - |x^H r|^2}, & \text{if } j \text{ is odd.} \end{cases}$$

Then $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ is such that $\triangle\mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: The proof is similar to Theorem 4.3.11.■

We mention that the results obtained above can be extended easily to polynomials having more general structures such as the case when the coefficient matrices are in Jordan and/or Lie algebras. Indeed, let M be a unitary matrix such that $M^T = M$ or $M^T = -M$. Consider the Jordan algebra $\mathbb{J} := \{A \in \mathbb{C}^{n \times n} : M^{-1}A^T M = A\}$ and the Lie algebra $\mathbb{L} := \{A \in \mathbb{C}^{n \times n} : M^{-1}A^T M = -A\}$ associated with the scalar product $(x, y) \mapsto y^T M x$. Consider a polynomial $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$, where A_j's are in $\mathbb{J}$ and/or in $\mathbb{L}$. Then the polynomial $M\mathrm{P}$

TH-789_04612301

given by $M\mathrm{P}(z) = \sum_{j=0}^{m} \lambda^j MA_j$ is either T-symmetric, T-skew-symmetric, T-even or T-odd. Hence replacing A_j and r by MA_j and Mr, respectively, in the above results, we obtain corresponding results for the polynomial P.

Similarly, when M is unitary and $M = M^H$ or $M = -M^H$, we consider the Jordan algebra $\mathbb{J} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^H M = A\}$ and the Lie algebra $\mathbb{L} := \{A \in \mathbb{C}^{n\times n} : M^{-1}A^H M = -A\}$ associated with the scalar product $(x, y) \mapsto y^H Mx$. Now, let $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$ be a polynomial where A_j's are in $\mathbb{J}$ and/or in $\mathbb{L}$. Then the polynomial $M\mathrm{P}(z) = \sum_{j=0}^{m} z^j MA_j$ is either H-Hermitian, H-skew-Hermitian, H-even or H-odd. Hence the results obtained above extend easily to the polynomial P by replacing A_j, r by MA_j, Mr, respectively. In particular, when $M := J$, where $J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$, the Jordan algebra $\mathbb{J}$ consists of skew-Hamiltonian matrices and the Lie algebra $\mathbb{L}$ consists of Hamiltonian matrices. So, for example, considering the polynomial $\mathrm{P}(z) := \sum_{j=0}^{m} z^j A_j$, where A_js are Hamiltonian when j is even and skew-Hamiltonian when j is odd, we see that the polynomial $J\mathrm{P}(z) = \sum_{j=0}^{m} z^j JA_j$ is H-even. Hence extending the results obtained for H-even polynomial to the case of P, we have the following.

Theorem 4.3.13. *Let $\mathbb{S}$ be the space of polynomials of the form $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$ where A_j is Hamiltonian when j is even, and A_j is skew-Hamiltonian when j is odd. Let $\mathrm{P} \in \mathbb{S}$. Then for $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, set $r := -\mathrm{P}(\lambda)x$. Then we have*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{\sqrt{2\|r\|_2^2 - |x^H Jr|^2}}{\|\Lambda_m\|_2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{2(\|r\|_2^2 - |x^H Jr|^2)}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R}. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda \in i\,\mathbb{R} \\ \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H Jr|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \in \mathbb{C} \setminus i\,\mathbb{R} \end{cases}$$

where $\widehat{r} = \begin{bmatrix} \Pi_e\,(\mathsf{Re}(\Lambda_m)^T) - (I - \Pi_e)(\mathsf{Im}(\Lambda_m)^T) \\ \Pi_e\,(\mathsf{Im}(\Lambda_m)^T) + (I - \Pi_e)(\mathsf{Re}(\Lambda_m)^T) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(x^H Jr) \\ \mathsf{im}(x^H Jr) \end{bmatrix}.$

4.4 Structured backward error and structured linearizations

Let $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ be given by $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$. The standard linearizations of P are the block-companion forms, the first companion form $C_1(z) = zX_1 + Y_1$ and the second companion form $C_2(z) = zX_2 + Y_2$, where $X_1 = X_2 = \mathrm{diag}(A_m, I_n, \ldots, I_n)$ and

$$Y_1 = \begin{bmatrix} A_{m-1} & A_{m-2} & \ldots & A_0 \\ -I_n & 0 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & -I_n & 0 \end{bmatrix}, \; Y_2 = \begin{bmatrix} A_{m-1} & -I_n & \ldots & 0 \\ A_{m-2} & 0 & \ldots & 0 \\ \vdots & \vdots & & -I_n \\ A_0 & 0 & \ldots & 0 \end{bmatrix}.$$

Both companion forms preserve algebraic and partial multiplicities of all finite eigenvalues of $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$. These also preserve the multiplicities of the eigenvalue ∞ and are called

TH-789_04612301

strong linearizations of P. Generalizing the concepts of companion forms, Mackey at al. [67] introduced two vector spaces of potential linearizations

$$\begin{aligned}\mathbb{L}_1(\mathrm{P}) &= \{\mathrm{L}(\lambda) : \mathrm{L}(\lambda).(\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda),\ v \in \mathbb{C}^m\}\\ \mathbb{L}_2(\mathrm{P}) &= \{\mathrm{L}(\lambda) : (\Lambda_{m-1}^T \otimes I_n).\mathrm{L}(\lambda) = \omega^T \otimes \mathrm{P}(\lambda),\ \omega \in \mathbb{C}^m\}\end{aligned}$$

of dimension $m(m-1)n^2 + m$, where $\Lambda_{m-1} = [\lambda^{m-1}, \lambda^{m-2}, \ldots, 1]^T$, $\otimes$ is the Kronecker product, v is called the right ansatz vector for $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ and ω is called the left ansatz vector for $\mathrm{L} \in \mathbb{L}_2(\mathrm{P})$. Observe that $C_1(\lambda) \in \mathbb{L}_1(\mathrm{P})$ with $v = e_1$ and $C_2(\lambda) \in \mathbb{L}_2(\mathrm{P})$ with $\omega = e_1$. It is also shown that $x \in \mathbb{C}^n$ is a right (resp. left) eigenvector of P corresponding to the eigenvalue λ if and only if $\Lambda_{m-1} \otimes x$(resp. $\overline{\Lambda_{m-1}} \otimes x$) is an eigenvector of $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$(resp. $\mathbb{L}_2(\mathrm{P})$) corresponding to the eigenvalue λ.

Since the linearization is unique up to the choice of right/left ansatz vector, these spaces are too large to obtain a potential linearization. Keeping this in mind, define the double ansatz space $\mathbb{DL}(\mathrm{P}) := \mathbb{L}_1(\mathrm{P}) \cap \mathbb{L}_2(\mathrm{P})$, see [64, 67]. It is also proved that for vectors $v = [v_1, v_2, \ldots, v_m]^T$ and $\omega = [\omega_1, \omega_2, \ldots, \omega_m]^T$, there exist an $mn \times mn$ matrix pencil $\mathrm{L}(\lambda) = \lambda X + Y \in \mathbb{DL}(\mathrm{P})$ that simultaneously satisfies

$$\mathrm{L}(\lambda) \cdot (\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda),\ (\Lambda_{m-1}^T \otimes I_n) \cdot \mathrm{L}(\lambda) = \omega^T \otimes \mathrm{P}(\lambda)$$

if and only if $v = \omega$, called the ansatz vector associated to L. Corresponding to an ansatz vector vector $v = [v_1, v_2, \ldots, v_m]^T \in \mathbb{C}^m$ associate the scalar polynomial $\mathsf{p}(x; v) = v_1 x^{m-1} + v_2 x^{m-2} + \ldots + v_{m-1}x + v_m$ referred to as the "v-polynomial" of the vector v, see [64, 67]. We adopt the convention that $\mathsf{p}(x; v)$ has a root at ∞ whenever $v_1 = 0$. Then $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ corresponding to an ansatz vector $v \in \mathbb{C}^m$ is a linearization of $\mathrm{P}(\lambda)$ if and only if no root of the v-polynomial is an eigenvalue of $\mathrm{P}(\lambda)$, see [64, 67]. Note that this statement includes ∞ as one of the possible roots of $p(x; \mathtt{v})$ or possible eigenvalue of $\mathrm{P}(\lambda)$.

Assume that $\mathrm{L}(\lambda) = \lambda X + Y \in \mathbb{L}_1(\mathrm{P})$ is a linearization of a polynomial $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j$ with respect to the normalized right ansatz vector $v \in \mathbb{C}^m$. Then if $x \in \mathbb{C}^n$ is a right eigenvector of P corresponding to an eigenvalue λ then $\Lambda_{m-1} \otimes x$ is a right eigenvector of L corresponding to the eigenvalue λ. Recall from Lemma 1.2.19 that for $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})/\mathbb{DL}(\mathrm{P})$ the following hold for any $x \in \mathbb{C}^n$.

$$\|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2 = \|v\|_2 \, \|\mathrm{P}(\lambda)x\|_2 \tag{4.7}$$

$$|(\Lambda_{m-1} \otimes x)^T \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)| = |\Lambda_{m-1}^T v| \, |x^T \mathrm{P}(\lambda)x| \tag{4.8}$$

$$|(\Lambda_{m-1} \otimes x)^H \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)| = |\Lambda_{m-1}^H v| \, |x^H \mathrm{P}(\lambda)x| \tag{4.9}$$

It is also straightforward to verify that

$$\frac{1}{\sqrt{2}} \le \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2 \, \|(\lambda, 1)\|_2} \le 1. \tag{4.10}$$

Assume that $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ and (λ, x) is an approximate eigenpair of P. Now we compare $\eta(\lambda, x, \mathrm{P})$ and $\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)$ where $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is a linearization of P corresponding to an ansatz vector v and $\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)$ denotes the unstructured backward error of the approximate eigenpair $(\lambda, \Lambda_{m-1} \otimes x)$ of L. It is needless to say that $\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)$

TH-789_04612301

can easily be obtained from (4.1) by setting $m = 1$. Without loss of generality we carry the analysis by taking unit ansatz vector, that is, $v \in \mathbb{C}^m$ is an ansatz vector with $\|v\|_2 = 1$.

Theorem 4.4.1. *Let* $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ *be a regular matrix polynomial and* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ *be a linearization of* P *corresponding to a right ansatz vector* $v \in \mathbb{C}^m$*. Assume that* (λ, x) *is an approximate eigenpair of* P*. Then we have*

$$\frac{1}{\sqrt{2}} \leq \frac{\eta(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1.$$

Proof: By (4.1), we have

$$\begin{aligned}\eta(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) &= \frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)\|_2}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(\lambda,\, 1)\|_2} = \frac{\|v\|_2\|\mathrm{P}(\lambda)x\|_2}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(\lambda,\, 1)\|_2} \\ &= \frac{\|v\|_2\|\Lambda_m\|_2\|x\|_2}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(\lambda,\, 1)\|_2}\eta(\lambda, x, \mathrm{P}) \\ &= \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2\,\|(\lambda,\, 1)\|_2}\eta(\lambda, x, \mathrm{P}).\end{aligned}$$

Now by (4.10) we obtain the desired result. ■

Its evident that a matrix polynomial can have infinitely many linearizations. For a given structured polynomial $\mathrm{P} \in \mathbb{S}$, it is necessary to detect a linearization L so that it preserves the spectral symmetry. Although, we have encountered three linear spaces of linearizations, namely $\mathbb{L}_1(\mathrm{P}), \mathbb{L}_2(\mathrm{P}), \mathbb{DL}(\mathrm{P})$, only $\mathbb{L}_1(\mathrm{P})$ preserves most of the structures which we consider in this chapter and it is easy to obtain the right eigenvector of P from that of $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$. In fact, for a T-symmetric polynomial P we have $\mathbb{DL}(\mathrm{P}) = \mathbb{L}_1(\mathrm{P})$. As shown in [40, 68], structured linearization imposes a restriction on the ansatz vector. To obtain a potential structured linearization we restrict the ansatz vector into a particular subset of $\mathbb{C}^m$ as shown by Mackey et al. [40, 68]. A list of structured linearizations and corresponding structure of ansatz vectors are given in the Table 4.2.

where $R = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$, $\qquad \Sigma = \mathrm{diag}\{(-1)^{m-1}, (-1)^{m-2}, \ldots, (-1)^0\}$.

Table 4.2 shows that except for T-symmetric and T-skew-symmetric polynomials, for all other structured polynomials P we could have two types of structured linearizations having the same spectral symmetry as that of P. A crucial task is to choose the best possible type of structured linearization for a given structured polynomial. In this section we provide a recipe for structured linearization L of a given $\mathrm{P} \in \mathbb{S}$.

Recall that $\eta(\lambda, x, \mathrm{P}) \leq \eta^{\mathbb{S}}_F(\lambda, x, \mathrm{P})$ and $\eta(\lambda, x, \mathrm{P}) \leq \eta^{\mathbb{S}}_2(\lambda, x, \mathrm{P})$. Hence for any structured linearization L we have $\eta(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) \leq \eta^{\mathbb{S}}_{F,2}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)$, where the ansatz vector v is of the form as described in Table 4.2. To make the presentation simple, unless otherwise stated, we write $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)$ for both $\eta^{\mathbb{S}}_F(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)$ and $\eta^{\mathbb{S}}_2(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)$, in the rest of the chapter.

Lemma 4.4.2. *Let* $\mathrm{P} \in \mathbb{S}$*. Let* $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ *be a structured linearization of* P *corresponding to*

TH-789_04612301

$\mathbb{S}$	Structured Linearization	ansatz vector
T-symm	T-symm	$v \in \mathbb{C}^m$
T-skew-symm	T-skew-symm	$v \in \mathbb{C}^m$
T-even	T-even	$\Sigma v = v$
	T-odd	$\Sigma v = -v$
T-odd	T-even	$\Sigma v = -v$
	T-odd	$\Sigma v = v$
H-Herm	H-Herm	$v \in \mathbb{R}^m$
	H-skew-Herm	$v \in i\mathbb{R}^m$
H-skew-Herm	H-Herm	$v \in i\mathbb{R}^m$
	H-skew-Herm	$v \in \mathbb{R}^m$
H-even	H-even	$\Sigma v = \overline{v}$
	T-odd	$\Sigma v = -\overline{v}$
H-odd	H-even	$\Sigma v = -\overline{v}$
	H-odd	$\Sigma v = \overline{v}$

Table 4.2: Table for the admissible ansatz vectors for structured polynomials.

an ansatz vector v. Then for both $|||\cdot||| \equiv |||\cdot|||_F$ and $|||\cdot||| \equiv |||\cdot|||_2$ we have

$$\frac{\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \geq \frac{1}{\sqrt{2}}.$$

Proof: By Theorem 4.4.1 we have $\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v) \geq \frac{1}{\sqrt{2}}\eta(\lambda, x, \mathrm{P})$. Therefore we obtain

$$\frac{\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \geq \frac{\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \geq \frac{1}{\sqrt{2}}.$$

■

Given an ansatz vector v, for the rest of the chapter, we set

$$\delta_v := \frac{\|\Lambda_{m-1}\|_2}{|\Lambda_{m-1}^* v|}, \; * \in \{T, H\}. \tag{4.11}$$

Obviously $\delta_v \geq 1$. Note that $\Lambda_{m-1}^* v \neq 0$ and hence $\delta_v < \infty$ when $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ is a linearization of P corresponding to the ansatz vector v.

For a T-symmetric matrix polynomial P, any ansatz vector v yields a T-symmetric linearization. Thus, we are free to choose v as followed by Table 4.2. Combined with Theorem 4.3.1, which shows that there is (almost) no difference between structured and unstructured backward errors, we have the following result.

Theorem 4.4.3. *Let $\mathbb{S}$ be the space of T-symmetric matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{L}_1(\mathrm{P}) = \mathbb{DL}(\mathrm{P})$ is the T-symmetric linearization of P with respect to an ansatz vector v. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ be such that $\|x\|_2 = 1$. Then we have*

1. $|||\cdot||| \equiv |||\cdot|||_F : \dfrac{\sqrt{2 - \delta_v^{-2}}}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$

TH-789_04612301

2. $\|\cdot\| \equiv \|\cdot\|_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$

Proof: First consider $\|\cdot\| \equiv \|\cdot\|_F$. By Theorem 4.3.1 we know that

$$\begin{aligned}
\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v) &= \frac{\sqrt{2\frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} - \frac{1}{\|(\Lambda_{m-1}\otimes x)\|_2^4}\,|(\Lambda_{m-1}\otimes x)^T\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)|^2}}{\|(\lambda,\, 1)\|_2} \\
&= \frac{\sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 - \frac{|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}|x^T\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\,\|(\lambda,\, 1)\|_2}, \text{ by (1.6), (1.7).}
\end{aligned}$$

Now applying $\sqrt{2-\delta_v^{-2}}\,\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 - \delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2}\,\|\mathrm{P}(\lambda)x\|_2$, (4.10) and (4.11) we obtain the desired result.

Next consider $\|\cdot\| \equiv \|\cdot\|_2$. Since structured backward error and unstructured backward error are same for spectral norm, the desired result follows by Theorem 4.4.1 and (4.10).■

Assuming $\|\cdot\| \equiv \|\cdot\|_F$ and $\|\cdot\| \equiv \|\cdot\|_2$ the results discussed above present growth of unstructured backward error of an eigenpair of $\mathrm{P} \in \mathbb{S}$ when P is linearized by a structured pencil in $\mathbb{L}_1(\mathrm{P}) = \mathbb{DL}(\mathrm{P})$. However, comparing $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)$ with $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ we have the following results for $\|\cdot\| \equiv \|\cdot\|_F$ and $\|\cdot\| \equiv \|\cdot\|_2$.

Theorem 4.4.4. *Let $\mathbb{S}$ be the space of T-symmetric matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{L}_1(\mathrm{P}) = \mathbb{DL}(\mathrm{P})$ is the T-symmetric linearization of* P *with respect to an ansatz vector v. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ be such that $\|x\|_2 = 1$. Then we have*

1. $\|\cdot\| \equiv \|\cdot\|_F : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} \leq \sqrt{2}$

2. $\|\cdot\| \equiv \|\cdot\|_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})} \leq 1.$

Proof: First consider $\|\cdot\| \equiv \|\cdot\|_F$. Then by Theorem 4.3.1 we have

$$\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} = \frac{\sqrt{2\|r\|_2^2 - \delta_v^{-2}|x^T r|^2}}{\sqrt{2\|r\|_2^2 - |x^T r|^2}} \cdot \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2\|(\lambda,\, 1)\|_2},\ r = -\mathrm{P}(\lambda)x.$$

Now using the inequality $0 < \delta_v^{-1} \leq 1$ by (4.11), and by (4.10) we have $\dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} \geq \frac{1}{\sqrt{2}}$. Further notice that

$$\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F(\lambda, x, \mathrm{P})}.$$

Now by Theorem 4.4.3 the desired result follows.

Next consider $\|\cdot\| \equiv \|\cdot\|_2$. Then by Theorem 4.3.1 we have

$$\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} = \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2\|(\lambda,\, 1)\|_2}.$$

Therefore by (4.10) we obtain the desired result.■

TH-789_04612301

Now we consider T-skew-symmetric polynomials. It follows from the Table 4.2 that any ansatz vector v yields a T-skew-symmetric linearization for a T-skew-symmetric matrix polynomial P. Combined with Theorem 4.3.1, which states that there is (almost) no difference between structured and unstructured backward errors, we have the following result.

Theorem 4.4.5. *Let $\mathbb{S}$ be the space of T-skew-symmetric matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is the T-skew-symmetric linearization of* P *with respect to the ansatz vector v. Let (λ, x) be an approximate eigenpair of* P*. Then we have*

1. $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_F : 1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$

2. $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1.$

Proof: First consider $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_F$. By Theorem 4.3.4 we have

$$\begin{aligned} \eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v) &= \sqrt{2} \frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2}{\|\Lambda_{m-1}\|_2 \, \|x\|_2 \, \|(1, \, \lambda)\|_2} \\ &= \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2 \, \|(\lambda, \, 1)\|_2} \eta(\lambda, x, \mathrm{P}) \end{aligned}$$

for any ansatz vector v. By (4.10) we obtain the desired result.

Since structured backward error and unstructured backward error are same for $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_2$, the desired result follows by Theorem 4.4.1 and (4.10).■

Further, comparing $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)$ with $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ we have the following results for $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_F$ and $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_2$.

Theorem 4.4.6. *Let $\mathbb{S}$ be the space of T-skew-symmetric matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{L}_1(\mathrm{P}) = \mathbb{DL}(\mathrm{P})$ is the T-skew-symmetric linearization of* P *with respect to an ansatz vector v. Let $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ be such that $\|x\|_2 = 1$. Then we have*

1. $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_F : 1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} \leq \sqrt{2}$

2. $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})} \leq 1.$

Proof: First consider $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_F$. Then by Theorem 4.3.4 it follows that

$$\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} = \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2 \|(\lambda, \, 1)\|_2}.$$

Hence by (4.10) we obtain the desired result.

Next consider $\|\!|\cdot|\!\| \equiv \|\!|\cdot|\!\|_2$. Then by Theorem 4.3.4 we have

$$\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})} = \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2 \|(\lambda, \, 1)\|_2}.$$

Therefore by (4.10)the desired result follows.■

Next consider T-even polynomials. Note that T-even polynomials can have both T-even and T-odd linearizations having the same eigen-symmetry as that of T-even polynomial.

TH-789_04612301

Theorem 4.4.7. *Let $\mathbb{S}$ be the space of T-even matrix polynomials. Assume that* $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j \in \mathbb{S}$. *Let* L_e *(resp.* L_o*) from* $\mathbb{L}_1(\mathrm{P})$ *be the T-even (resp. T-odd) linearization of* P *with respect to the ansatz vector* $\Sigma v = v$ *(resp.* $\Sigma v = -v$*). If* (λ, x) *is an approximate eigenpair of* P *then we have the following.*

1. $|||\cdot||| \equiv |||\cdot|||_F$: *If* $|\lambda| \leq 1$ *then* $\dfrac{\sqrt{2+(|\lambda|^2-1)\delta_v^{-2}}}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ *and*

   $1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2+(|\lambda|^{-2}-1)\delta_v^{-2}}.$

   If $|\lambda| \geq 1$ *then* $1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2+(|\lambda|^2-1)\delta_v^{-2}}$ *and*

   $\dfrac{\sqrt{2+(|\lambda|^{-2}-1)\delta_v^{-2}}}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$

2. $|||\cdot||| \equiv |||\cdot|||_2$: $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{1+|\lambda|^2\delta_v^{-2}}$

   $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{1+|\lambda|^{-2}\delta_v^{-2}}$ *if* $\lambda \neq 0$,

   $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$, *if* $\lambda = 0$.

Proof: First consider the T-even linearization L_e of $\mathrm{P} \in \mathbb{S}$. By Theorem 4.3.6 we know that

$$\begin{aligned}\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v) &= \frac{\sqrt{2\frac{\|\mathrm{L}_e(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} + \frac{(|\lambda|^2-1)}{\|(\Lambda_{m-1}\otimes x)\|_2^4}|(\Lambda_{m-1}\otimes x)^T\mathrm{L}_e(\lambda)(\Lambda_{m-1}\otimes x)|^2}}{\|(\lambda,\, 1)\|_2} \\ &= \frac{\sqrt{2\,\|\mathrm{P}(\lambda)x\|_2^2 + \frac{(|\lambda|^2-1)|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}\,|x^T\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\,\|(\lambda,\, 1)\|_2}.\end{aligned}$$

For $|\lambda| \leq 1$, we have $\sqrt{2+(|\lambda|^2-1)\delta_v^{-2}}\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 + (|\lambda|^2-1)\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2}\,\|\mathrm{P}(\lambda)x\|_2$ and for $|\lambda| > 1$ we have

$$\sqrt{2}\,\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\,\|\mathrm{P}(\lambda)x\|_2^2 + (|\lambda|^2-1)\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2+(|\lambda|^2-1)\delta_v^{-2}}\|\mathrm{P}(\lambda)x\|_2.$$

Hence by (4.10) we obtain the desired results for $|||\cdot||| \equiv |||\cdot|||_F$.

Next, consider $|||\cdot||| \equiv |||\cdot|||_2$. Then by Theorem 4.3.6 we have

$$\begin{aligned}\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v) &= \frac{\sqrt{\frac{\|\mathrm{L}_e(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} + \frac{|\lambda|^2}{\|(\Lambda_{m-1}\otimes x)\|_2^4}|(\Lambda_{m-1}\otimes x)^T\mathrm{L}_e(\lambda)(\Lambda_{m-1}\otimes x)|^2}}{\|(\lambda,\, 1)\|_2} \\ &= \frac{\sqrt{\|\mathrm{P}(\lambda)x\|_2^2 + \frac{|\lambda|^2\,|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}\,|x^T\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\,\|(\lambda,\, 1)\|_2}.\end{aligned}$$

Notice that $\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{\|\mathrm{P}(\lambda)x\|_2^2 + |\lambda|^2\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2+|\lambda|^2\delta_v^{-2}}\,\|\mathrm{P}(\lambda)x\|_2$ for any $\lambda \in \mathbb{C}$. Hence by (4.10) we obtain the desired result.

Next assume that L_o is the T-odd linearization of $\mathrm{P} \in \mathbb{S}$. By Theorem 4.3.8 we have the

TH-789_04612301

following for $\|\cdot\| \equiv \|\cdot\|_F$:

$$\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v) = \begin{cases} \dfrac{\sqrt{2\frac{\|\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} + (\frac{1}{|\lambda|^2}-1)\frac{1}{\|(\Lambda_{m-1}\otimes x)\|_2^4}|(\Lambda_{m-1}\otimes x)^T\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)|^2}}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(1,\lambda)\|_2}, & \\ & \textit{if } \lambda \neq 0. \\ \sqrt{2}\,\dfrac{\|\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)\|_2}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(1,\lambda)\|_2}, & \textit{if } \lambda = 0. \end{cases}$$

$$= \begin{cases} \dfrac{\sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 + (\frac{1}{|\lambda|^2}-1)\frac{|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}|x^T\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\|x\|_2\,\|(1,\lambda)\|_2}, & \text{if } \lambda \neq 0. \\ \sqrt{2}\,\dfrac{\|\mathrm{P}(\lambda)x\|_2}{\|\Lambda_{m-1}\|_2\|x\|_2\,\|(1,\lambda)\|_2}, & \text{if } \lambda = 0. \end{cases}$$

Let $\lambda \neq 0$. Then for $|\lambda| \leq 1$, we have $\sqrt{2}\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 + (|\lambda|^{-2}-1)\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2 + (|\lambda|^{-2}-1)\delta_v^{-2}}\|\mathrm{P}(\lambda)x\|_2$ and for $|\lambda| < 1$ we have

$$\sqrt{2 + (|\lambda|^{-2}-1)\delta_v^{-2}}\,\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 + (|\lambda|^{-2}-1)\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{2}\|\mathrm{P}(\lambda)x\|_2.$$

Hence by (4.10), we obtain the desired result.

Now consider $\|\cdot\| \equiv \|\cdot\|_2$. Then by Theorem 4.3.8 we have

$$\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v) = \begin{cases} \dfrac{\sqrt{\frac{\|\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} + \frac{1}{|\lambda|^2\,\|(\Lambda_{m-1}\otimes x)\|_2^4}|(\Lambda_{m-1}\otimes x)^T\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)|^2}}{\|(1,\lambda)\|_2}, & \\ & \textit{if } \lambda \neq 0. \\ \dfrac{\|\mathrm{L}_o(\lambda)(\Lambda_{m-1}\otimes x)\|_2}{\|(\Lambda_{m-1}\otimes x)\|_2\,\|(1,\lambda)\|_2}, & \textit{if } \lambda = 0 \end{cases}$$

$$= \begin{cases} \dfrac{\sqrt{\|\mathrm{P}(\lambda)x\|_2^2 + \frac{|\Lambda_{m-1}^T v|^2}{|\lambda|^2\,\|\Lambda_{m-1}\|_2^2}|x^T\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\|(1,\lambda)\|_2}, & \text{if } \lambda \neq 0. \\ \dfrac{\|\mathrm{P}(\lambda)x\|_2}{\|\Lambda_{m-1}\|_2\,\|(1,\lambda)\|_2}, & \text{if } \lambda = 0 \end{cases}$$

For $\lambda \neq 0$ notice that $\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{\|\mathrm{P}(\lambda)x\|_2^2 + |\lambda|^{-2}\delta_v^{-2}|x^T\mathrm{P}(\lambda)x|^2} \leq \sqrt{1 + |\lambda|^{-2}\delta_v^{-2}}\|\mathrm{P}(\lambda)x\|_2$. Hence by (4.10) the desired result follows.■

Corollary 4.4.8. *For $\|\cdot\| \equiv \|\cdot\|_F$, using that fact that $\delta_v^{-1} > 0$ for any ansatz vector v, we have $1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ when $|\lambda| \leq 1$ and $1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ when $|\lambda| \geq 1$. For $\|\cdot\| \equiv \|\cdot\|_2$, using the fact that $|\delta_v^{-1}| \leq 1$ for any ansatz vector v, we have $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_e; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ when $|\lambda| \leq 1$, and $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_o; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ when $|\lambda| \geq 1$.*

Now notice that the bound in Theorem 4.4.7 is true also for T-odd polynomials but with the roles of T-even and T-odd exchanged.

Remark 4.4.9. *Let* P *be a T-even or T-odd polynomial. Then if $|\lambda| \leq 1$ pick the T-even linearization and for $|\lambda| > 1$ pick the T-odd linearization. The corrolarry 4.4.8 also ensure that the backward error is magnified at most by a factor of $\sqrt{2}$ due to the structured linearization process.*

TH-789_04612301

Next we consider H-Hermitian polynomials. Note that H-Hermitian polynomial can have both H-Hermitian and H-skew-Hermitian linearizations preserving the eigen-symmetry of the original polynomial.

Theorem 4.4.10. *Let $\mathbb{S}$ be the space of H-Hermitian matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ is the H-Hermitian/ H-skew-Hermitian linearization of P with respect to the ansatz vector v. If (λ, x) is an approximate right eigenpair of P then we have the following:*

1. $|||\cdot||| \equiv |||\cdot|||_F : \dfrac{\sqrt{2-\delta_v^{-2}}}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ *if* $\lambda \in \mathbb{R}$,

   $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{\tau^2\delta_v^{-2}\|(1,\,\lambda)\|_2^2 + 2}$, *if* $\lambda \in \mathbb{C}\setminus\mathbb{R}$

2. $|||\cdot||| \equiv |||\cdot|||_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$ *if* $\lambda \in \mathbb{R}$,

   $\dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{\tau^2\delta_v^{-2}\|(1,\,\lambda)\|_2^2 + 1}$, *if* $\lambda \in \mathbb{C}\setminus\mathbb{R}$

where $\tau = \left\| \begin{bmatrix} 1 & \mathsf{re}(\lambda) \\ 0 & \mathsf{im}(\lambda) \end{bmatrix}^{\dagger} \right\|_2$ *for H-Hermitian linearization and* $\tau = \left\| \begin{bmatrix} 0 & -\mathsf{im}(\lambda) \\ 1 & \mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2$ *for H-skew-Hermitian linearization.*

Proof: First we consider the H-Hermitian linearization. Then the ansatz vector $v \in \mathbb{R}^m$. Assume that $\lambda \in \mathbb{R}$. Then by Theorem 4.3.9 we have

$$\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) = \frac{\sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 - \delta_v^{-2}|x^H\mathrm{P}(\lambda)x|^2}}{\|\Lambda_{m-1}\|_2\,\|(1,\lambda)\|_2},\ \eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) = \frac{\|\mathrm{P}(\lambda)x\|_2}{\|\Lambda_{m-1}\|_2\,\|(\lambda,\,1)\|_2}.$$

Now applying $\sqrt{2-\delta_v^{-2}}\|\mathrm{P}(\lambda)x\|_2 \leq \sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 - \delta_v^{-2}|x^H\mathrm{P}(\lambda)x|^2} \leq \sqrt{2}\|\mathrm{P}(\lambda)x\|_2$ and (4.10) we obtain the desired result.

Next consider $\lambda \in \mathbb{C}\setminus\mathbb{R}$. Then for $|||\cdot||| \equiv |||\cdot|||_F$, by Theorem 4.3.9 we have

$$\begin{aligned}\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) &= \frac{1}{\|\Lambda_{m-1}\|_2\|x\|_2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^2} + \frac{2\|\mathrm{P}(\lambda)x\|_2^2 - 2\frac{|\Lambda_{m-1}^H v|^2}{\|\Lambda_{m-1}\|_2^2}|x^H\mathrm{P}(\lambda)x|^2}{\|(1,\,\lambda)\|_2^2}} \\ &\leq \frac{1}{\|\Lambda_{m-1}\|_2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^2} + \frac{2\|r\|_2^2}{\|(1,\,\lambda)\|_2^2}},\ r := -\mathrm{P}(\lambda)x\end{aligned}$$

where $\widehat{r} = \begin{bmatrix} 1 & \mathsf{re}\lambda \\ 0 & \mathsf{im}\lambda \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(\Lambda_{m-1}^H v x^H\mathrm{P}(\lambda)x) \\ \mathsf{im}(\Lambda_{m-1}^H v x^H\mathrm{P}(\lambda)x) \end{bmatrix}$. For $|||\cdot||| \equiv |||\cdot|||_2$, by Theorem 4.3.9 we have the following:

$$\begin{aligned}\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) &= \frac{1}{\|\Lambda_{m-1}\|_2\|x\|_2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^2} + \frac{\|\mathrm{P}(\lambda)x\|_2^2 - \frac{|\Lambda_{m-1}^H v|^2}{\|\Lambda_{m-1}\|_2^2}|x^H\mathrm{P}(\lambda)x|^2}{\|(1,\,\lambda)\|_2^2}} \\ &\leq \frac{1}{\|\Lambda_{m-1}\|_2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^2} + \frac{\|\mathrm{P}(\lambda)x\|_2^2}{\|(1,\,\lambda)\|_2^2}}\end{aligned}$$

TH-789_04612301

Now notice that $\|\widehat{r}\|_2 \leq \tau |\Lambda_{m-1}^H v| \|r\|_2$, where $\tau = \left\| \begin{bmatrix} 1 & \mathsf{re}\lambda \\ 0 & \mathsf{im}\lambda \end{bmatrix}^{\dagger} \right\|_2$.

Therefore for $|||\cdot||| \equiv |||\cdot|||_F$ we have $\frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2} \sqrt{\tau^2 \delta_v^{-2} + 2\|(1,\, \lambda)\|_2^{-2}}$, and $\frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2} \sqrt{\tau^2 \delta_v^{-2} + \|(1,\, \lambda)\|_2^{-2}}$. Hence by Lemma 4.4.2 the results follows.

Next assume that L is the H-skew-Hermitian linearization of $\mathrm{P} \in \mathbb{S}$ associated with the ansatz vector $v \in i\mathbb{R}^m$. Since the structured backward error expression of any approximate eigenpair is same for H-Hermitian and H-skew-Hermitian linear polynomial, we obtain the same bound.■

Note that, for $\lambda \in \mathbb{R}$, it does not really make any difference in the magnified backward error due to H-Hermitian or H-skew-Hermitian linearization. Further for $\lambda \in \mathbb{C} \setminus \mathbb{R}$ two cases arise. Let us redefine $\tau := \tau_h$ for H-Hermitian linearization and $\tau := \tau_{sh}$ for H-skew-Hermitian linearization. Then from the bounds of Theorem 4.4.10 we conclude the following. For a given (λ, x), if $\tau_h \leq \tau_{sh}$ then choose H-Hermitian linearization and H-skew-Hermitian linearization otherwise.

Next consider H-even polynomials. Note that H-even polynomials can have both H-even and H-odd types of linearizations having the same eigen-symmetry $(\lambda, -\overline{\lambda})$. The main issue here is to decide which type of linearization is to be chosen to minimize the backward error.

Theorem 4.4.11. *Let $\mathbb{S}$ be the space of H-even polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is the H-even/H-odd linearization of P with respect to the ansatz vector v. If $(\lambda,\, x)$ is an approximate eigenpair of P then we have the following:*

1. $|||\cdot||| \equiv |||\cdot|||_F$: $\frac{\sqrt{2 - \delta_v^{-2}}}{\sqrt{2}} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ *if* $\lambda \in i\mathbb{R}$,

   $\frac{1}{\sqrt{2}} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{\tau^2 \delta_v^{-2} \|(1,\, \lambda)\|_2^2 + 2}$, *if* $\lambda \in \mathbb{C} \setminus i\mathbb{R}$

2. $|||\cdot||| \equiv |||\cdot|||_2$: $\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$ *if* $\lambda \in i\mathbb{R}$,

   $\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{\tau^2 \delta_v^{-2} \|(1,\, \lambda)\|_2^2 + 1}$, *if* $\lambda \in \mathbb{C} \setminus i\mathbb{R}$

where $\tau = \left\| \begin{bmatrix} 1 & -\mathsf{im}(\lambda) \\ 0 & \mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2$ *for H-even linearization and* $\tau = \left\| \begin{bmatrix} 0 & \mathsf{re}(\lambda) \\ 1 & \mathsf{im}(\lambda) \end{bmatrix}^{\dagger} \right\|_2$ *for H-odd linearization.*

Proof: The proof is similar as that of Theorem 4.4.10 and hence omitted.■

The similar bound can be obtained if we replace H-even polynomials by H-odd polynomials in Theorem 4.4.11.

Observe that, for $\lambda \in i\mathbb{R}$, it does not really make any difference in the magnified backward error due to H-even or H-odd linearization. Further for $\lambda \in \mathbb{C} \setminus i\mathbb{R}$ two cases arise. Let us redefine $\tau := \tau_e$ for H-even linearization and $\tau := \tau_o$ for H-odd linearization. Then from the bounds that have been obtained in Theorem 4.4.11, we conclude the following. For a given (λ, x) and $\mathrm{P} \in \mathbb{S}$, pick an H-even linearization when $\tau_e \leq \tau_o$ and pick H-odd linearization otherwise.

TH-789_04612301

structured polynomial	eigenvalue of interest	choice of linearization	$\beta^{\mathbb{S}} := \eta^{S}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)/\eta(\lambda, x, \mathrm{P})$ $\vert\!\vert\!\vert\cdot\vert\!\vert\!\vert \equiv \vert\!\vert\!\vert\cdot\vert\!\vert\!\vert_F$	$\beta^{\mathbb{S}} := \eta^{S}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)/\eta(\lambda, x, \mathrm{P})$ $\vert\!\vert\!\vert\cdot\vert\!\vert\!\vert \equiv \vert\!\vert\!\vert\cdot\vert\!\vert\!\vert_2$
T-symmetric	$\lambda \in \mathbb{C}$	T-symmetric	$\frac{\sqrt{2-\delta_v^{-2}}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 1$
T-skew-symmetric	$\lambda \in \mathbb{C}$	T-skew-symmetric	$1 \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 1$
T-even/T-odd	$\vert\lambda\vert \leq 1$	T-even	$\frac{\sqrt{2+(\vert\lambda\vert^2-1)\delta_v^{-2}}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$
	$\vert\lambda\vert \geq 1$	T-odd	$\frac{\sqrt{2+(\vert\lambda\vert^{-2}-1)\delta_v^{-2}}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	
H-Hermitian/ H-skew-Hermitian	$\lambda \in \mathbb{R}$	H-Hermitian/ H-skew-Hermitian	$\frac{\sqrt{2-\delta_v^{-2}}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 1$
	$\lambda \in \mathbb{C}\setminus\mathbb{R}$	H-Hermitian	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{2+\tau_h^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1+\tau_h^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$
		H-skew-Hermitian	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{2+\tau_{sh}^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1+\tau_h^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$
H-even/H-odd	$\lambda \in i\mathbb{R}$	H-even/H-odd	$\frac{\sqrt{2-\delta^2}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 1$
	$\lambda \in \mathbb{C}\setminus i\mathbb{R}$	H-even	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{2+\tau_e^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1+\tau_e^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$
		H-odd	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{2+\tau_o^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1+\tau_o^2\delta_v^{-2}\\|(1,\lambda)\\|_2^2}$

Table 4.3: The choice of structured linearizations, where $\delta_v = \|\Lambda_{m-1}\|_2/|\Lambda_{m-1}^{*}v|$, $* \in \{T, H\}$ and v is the ansatz vector.

$$\text{Here } r = -\mathrm{P}(\lambda)x, \tau_h = \left\| \begin{bmatrix} 1 & \mathsf{re}(\lambda) \\ 0 & \mathsf{im}(\lambda) \end{bmatrix}^{\dagger} \right\|_2, \tau_{sh} = \left\| \begin{bmatrix} 0 & -\mathsf{im}(\lambda) \\ 1 & \mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2, \tau_e = \left\| \begin{bmatrix} 1 & -\mathsf{im}(\lambda) \\ 0 & \mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2, \tau_o = \left\| \begin{bmatrix} 0 & \mathsf{re}(\lambda) \\ 1 & \mathsf{im}(\lambda) \end{bmatrix}^{\dagger} \right\|_2.$$

TH-789_04612301

4.5 Structured pseudospectra of structured matrix polynomials

Let P be a regular polynomial. For $\lambda \in \mathbb{C}$, the backward error of λ as an approximate eigenvalue of P is given by

$$\eta(\lambda, \mathrm{P}) := \min\{\eta(\lambda, x, \mathrm{L}) : x \in \mathbb{C}^n \text{ and } \|x\|_2 = 1\}.$$

Since $\eta(\lambda, x, \mathrm{P}) = \|r\|/\|\Lambda_m\|_2$, it follows that for the spectral as well as for the Frobenius norms on $\mathbb{C}^{n\times n}$, we have

$$\eta(\lambda, \mathrm{P}) := \frac{\sigma_{\min}(\mathrm{P}(\lambda))}{\|\Lambda_m\|_2}.$$

Similarly, we define structured backward error of an approximate eigenvalue λ of $\mathrm{P} \in \mathbb{S}$ by

$$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) := \min\{\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) : x \in \mathbb{C}^n \text{ and } \|x\|_2 = 1\}$$

for both $|\!|\!|\cdot|\!|\!| \equiv |\!|\!|\cdot|\!|\!|_F$ and $|\!|\!|\cdot|\!|\!| \equiv |\!|\!|\cdot|\!|\!|_2$. In this section, we make an attempt to determine $\eta^{\mathbb{S}}(\lambda, \mathrm{P})$. Note that backward errors of approximate eigenvalues and pseudospectra of a polynomial are closely related. For $\epsilon > 0$, the unstructured ϵ-pseudospectrum of P, denoted by $\sigma_\epsilon(\mathrm{P})$,

$$\sigma_\epsilon(\mathrm{P}) = \bigcup_{|\!|\!|\triangle \mathrm{P}|\!|\!| \leq \epsilon} \sigma(\mathrm{P} + \triangle \mathrm{P}).$$

See [1, 2] for more on pseudospectra of matrix polynomials. Obviously, we have $\sigma_\epsilon(\mathrm{P}) = \{z \in \mathbb{C} : \eta(z, \mathrm{P}) \leq \epsilon\}$, assuming, for simplicity, that $\infty \notin \sigma_\epsilon(\mathrm{P})$. For the sake of simplicity, for rest of this section, we make an implicit assumption that $\infty \notin \sigma_\epsilon(\mathrm{P})$. We observe the following.

- Since $\eta(\lambda, \mathrm{P})$ is the same for the spectral norm and the Frobenius norm on $\mathbb{C}^{n\times n}$, it follows that $\sigma_\epsilon(\mathrm{P})$ is the same for the spectral and the Frobenius norms.

Similarly, when $\mathrm{P} \in \mathbb{S}$, we define the structured ϵ-pseudospectrum of P, denoted by $\sigma_\epsilon^{\mathbb{S}}(\mathrm{P})$, by

$$\sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) := \bigcup_{|\!|\!|\triangle \mathrm{P}|\!|\!| \leq \epsilon} \{\sigma(\mathrm{P} + \triangle \mathrm{P}) : \triangle \mathrm{P} \in \mathbb{S}\}.$$

Then it follows that $\sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) = \{z \in \mathbb{C} : \eta^{\mathbb{S}}(\lambda, \mathrm{P}) \leq \epsilon\}$.

Theorem 4.5.1. *Let $\mathbb{S} \in \{T$-symmetric, T-skew-symmetric$\}$ and $\mathrm{P} \in \mathbb{S}$. Let $\lambda \in \mathbb{C}$. Then for $|\!|\!|\cdot|\!|\!| \equiv |\!|\!|\cdot|\!|\!|_2$ we have*

$$\eta_2^{\mathbb{S}}(\lambda, \mathrm{P}) = \eta(\lambda, \mathrm{P}) \text{ and } \sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) = \sigma_\epsilon(\mathrm{P}).$$

Also for $|\!|\!|\cdot|\!|\!| \equiv |\!|\!|\cdot|\!|\!|_F$ we have

$$\eta_F^{\mathbb{S}}(\lambda, \mathrm{P}) = \sqrt{2}\,\eta(\lambda, \mathrm{P}), \text{ and } \sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) = \sigma_{\epsilon/\sqrt{2}}(\mathrm{P})$$

when P is T-skew-symmetric and

$$\eta_F^{\mathbb{S}}(\lambda, \mathrm{P}) = \eta(\lambda, \mathrm{P}) \text{ and } \sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) = \sigma_\epsilon(\mathrm{P})$$

when P is T-symmetric.

TH-789_04612301

Proof: For the spectral norm, by Theorem 4.3.1, we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \eta(\lambda, x, \mathrm{P})$ for all x. Consequently, we have $\eta^{\mathbb{S}}(\lambda, \mathrm{P}) = \eta(\lambda, \mathrm{P})$. Hence the result follows.

For the Frobenius norm, the result follows from Theorem 4.3.4 when P is T-skew-symmetric. So, suppose that P is T-symmetric. Then $\mathrm{P}(\lambda) \in \mathbb{C}^{n\times n}$ is symmetric. Consider the Takagi factorization $P(\lambda) = U\Sigma U^T$, where U is unitary and Σ is a diagonal matrix containing singular values of $\mathrm{P}(\lambda)$ (appear in descending order). Set $\sigma := \Sigma(n, n)$ and $u := U(:, n)$. Then we have $\mathrm{P}(\lambda)\overline{u} = \sigma u$. Now define

$$\triangle A_j := -\frac{\overline{\lambda}^j \sigma\, uu^T}{\|\Lambda_m\|_2^2},$$

and consider the polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Then $\triangle\mathrm{P}$ is T-symmetric and $\mathrm{P}(\lambda)\overline{u} + \triangle\mathrm{P}(\lambda)\overline{u} = 0$. Notice that, for the $\|\cdot\| \equiv \|\cdot\|_2$ and $\|\cdot\| \equiv \|\cdot\|_F$ on $\mathbb{C}^{n\times n}$, we have

$$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) \leq \|\triangle\mathrm{P}\| = \frac{\sigma}{\|\Lambda_m\|_2} = \eta(\lambda, \mathrm{P}) \text{ and hence } \sigma_\epsilon(\mathrm{P}) = \sigma_\epsilon^{\mathbb{S}}(\mathrm{P}).$$

This completes the proof. ■

When P is T-symmetric, the above proof shows how to construct a T-symmetric pencil $\triangle\mathrm{P}$ such that $\lambda \in \Lambda_m(\mathrm{P} + \triangle\mathrm{P})$ and $\|\triangle\mathrm{P}\| = \eta^{\mathbb{S}}(\lambda, \mathrm{P})$. When P is T-skew-symmetric, using Takagi factorization of the complex skew-symmetric matrix $\mathrm{P}(\lambda)$, one can construct a T-skew-symmetric pencil $\triangle\mathrm{P}$ such that $\lambda \in \Lambda_m(\mathrm{P} + \triangle\mathrm{P})$ and $\|\triangle\mathrm{P}\| = \eta^{\mathbb{S}}(\lambda, \mathrm{P})$. Indeed, consider the Takagi factorization

$$\mathrm{P}(\lambda) = U\mathrm{diag}(d_1, \cdots, d_m)U^T,$$

where U is unitary, $d_j := \begin{bmatrix} 0 & s_j \\ -s_j & 0 \end{bmatrix}$, $s_j \in \mathbb{C}$ is nonzero and $|s_j|$ are singular values of $\mathrm{P}(\lambda)$. Here the blocks d_j appear in descending order of magnitude of $|s_j|$. Note that $\mathrm{P}(\lambda)\overline{U} = U\mathrm{diag}(d_1, \cdots, d_m)$. Let $u := U(:, n-1 : n)$. Then $\mathrm{P}(\lambda)\overline{u} = ud_m = ud_m u^T \overline{u}$. Now define

$$\triangle A_j := -\frac{\overline{\lambda^j}\, ud_m u^T}{\|v_\lambda\|_2^2}$$

and consider the pencil $\triangle\mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$. Then $\triangle\mathrm{P}$ is T-skew-symmetric and $\mathrm{P}(\lambda)\overline{u} + \triangle\mathrm{P}(\lambda)\overline{u} = 0$. For the spectral norm on $\mathbb{C}^{n\times n}$, we have

$$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) = \|\triangle\mathrm{P}\| = \frac{\sigma_{\min}(\mathrm{P}(\lambda))}{\|\Lambda_m\|_2} = \eta(\lambda, \mathrm{P})$$

and for the Frobenius norm on $\mathbb{C}^{n\times n}$, we have

$$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) = \|\triangle\mathrm{P}\| = \sqrt{2}\,\frac{\sigma_{\min}(\mathrm{P}(\lambda))}{\|\Lambda_m\|_2} = \sqrt{2}\,\eta(\lambda, \mathrm{P}).$$

We denote the unit circle in $\mathbb{C}$ by $\mathbb{T}$, that $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. Then for the T-even or T-odd polynomials we have the following result.

Theorem 4.5.2. *Let* $\mathbb{S} \in \{T\text{-even}, T\text{-odd}\}$ *and* $\mathrm{P} \in \mathbb{S}$ *and* m *is odd. Let* $\lambda \in \mathbb{T}$. *Then for the*

TH-789_04612301

Frobenius norm on $\mathbb{C}^{n\times n}$*, we have*

$$\eta^{\mathbb{S}}(\lambda,\mathrm{P}) = \sqrt{2}\,\eta(\lambda,\mathrm{P}) \text{ and } \sigma_\epsilon^{\mathbb{S}}(\mathrm{P})\cap\mathbb{T} = \sigma_{\epsilon/\sqrt{2}}(\mathrm{P})\cap\mathbb{T}.$$

Proof: Let $\lambda\in\mathbb{T}$. Then by Theorem 4.3.6 and Theorem 4.3.8, we have

$$\eta^{\mathbb{S}}(\lambda,x,\mathrm{P}) = \frac{\sqrt{2}\,\|\mathrm{P}(\lambda)x\|_2}{\|\Lambda_m\|_2}$$

for all x such that $\|x\|_2 = 1$. Hence taking minimum over $\|x\|_2 = 1$, we obtain the desired results. ■

Theorem 4.5.3. *Let* $\mathbb{S}\in\{H\text{-Hermitian}, H\text{-skew-Hermitian}\}$ *and* $\mathrm{P}\in\mathbb{S}$*. Let* $\lambda\in\mathbb{R}$*. Then for the spectral and the Frobenius norms on* $\mathbb{C}^{n\times n}$*, we have* $\eta^{\mathbb{S}}(\lambda,\mathrm{P}) = \eta(\lambda,\mathrm{P})$ *and hence*

$$\sigma_\epsilon^{\mathbb{S}}(\mathrm{P})\cap\mathbb{R} = \sigma_\epsilon(\mathrm{P})\cap\mathbb{R}.$$

Proof: Note that $\mathrm{P}(\lambda)$ is either Hermitian or skew-Hermitian. Let (μ,u) be an eigenpair of the matrix $\mathrm{P}(\lambda)$ such that $|\mu| = \sigma_{\min}(\mathrm{P}(\lambda))$ and $u^Hu = 1$. Then $\mathrm{P}(\lambda)u = \mu u$. Define

$$\triangle A_j := -\frac{\lambda^j\,\mu\,uu^H}{\|\Lambda_m\|_2^2}$$

and consider the pencil $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j\triangle A_j$. Then $\triangle\mathrm{P}\in\mathbb{S}$ and $\lambda\in\Lambda_m(\mathrm{P}+\triangle\mathrm{P})$. Further, for the spectral and the Frobenius norms, we have $\|\triangle\mathrm{P}\| = \dfrac{\sigma_{\min}(\mathrm{P}(\lambda))}{\|\Lambda_m\|_2}$. Hence the result follows. ■

Theorem 4.5.4. *Let* $\mathbb{S}\in\{H\text{-even}, H\text{-odd}\}$ *and* $\mathrm{P}\in\mathbb{S}$*. Let* $\lambda\in i\mathbb{R}$*. Then for the spectral and the Frobenius norms on* $\mathbb{C}^{n\times n}$*, we have* $\eta^{\mathbb{S}}(\lambda,\mathrm{P}) = \eta(\lambda,\mathrm{P})$ *and hence*

$$\sigma_\epsilon^{\mathbb{S}}(\mathrm{P})\cap i\,\mathbb{R} = \sigma_\epsilon(\mathrm{P})\cap i\,\mathbb{R}.$$

Proof: Note for $\lambda\in i\,\mathbb{R}$, then the matrix $\mathrm{P}(\lambda)$ is again either is Hermitian or skew-Hermitian. Hence the result follows from the proof of Theorem 4.5.3. ■

We mention that the above results can be easily extended to the case of general structured polynomials where the coefficients matrices are elements of Jordan and/or Lie algebras.

TH-789_04612301

Chapter 5

Backward errors and linearizations for palindromic matrix polynomials

This chapter is devoted to the backward perturbation analysis of palindromic and anti-palindromic polynomials. We derive structured backward error of approximate eigenpair of these polynomials and characterize the minimal structured perturbations that achieve it. Following similar approach employed for structured polynomials in chapter 4, we show that there always exists "good" palindromic /anti-palindromic linearization for a palindromic /anti-palindromic polynomial.

5.1 Introduction

In this chapter we undertake a detailed backward perturbation analysis of palindromic and anti-palindromic polynomials. These polynomials arise mainly in the study of rail traffic noise caused by high speed trains, see [44, 64, 68, 82]. A matrix polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is called a palindromic matrix polynomial if $\mathrm{P}^*(z) = z^m \mathrm{P}(1/z)$ for all $z \in \mathbb{C}\setminus\{0\}$ and an anti-palindromic polynomial if $\mathrm{P}^*(z) = z^m \mathrm{P}(-1/z)$ for all $z \in \mathbb{C}\setminus\{0\}$, where $*$ is the transpose or conjugate transpose of a matrix and

$$\mathrm{P}^*(z) = \sum_{j=0}^{m} z^j A_j^* \text{ whenever } \mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j.$$

In this chapter we consider only the regular matrix polynomials. We denote the space of regular palindromic or anti-palindromic polynomials by $\mathbb{S}$. It is shown in [64, 68] that eigenvalues of palindromic polynomials possess certain spectral symmetry.

It is well known that linearization is used to solve a given polynomial eigenvalue problem. As discussed in Chapter 4, an arbitrary linearization destroys the symmetry in the computed eigenvalues. Therefore, the first step towards solving a palindromic eigenvalue problem is to convert the given polynomial $\mathrm{P} \in \mathbb{S}$ into a linear polynomial L, call it a structured linearization which has the same eigensymmetry as that of P. It is shown in [68] that a palindromic or anti-palindromic polynomial has infinitely many palindromic/ and anti-palindromic linearizations

TH-789_04612301

which preserve the eigen-symmetry.

This gives rise to a problem of choosing a "good" linearization among them. Aiming to find a "good" structured linearization we follow similar procedure as discussed in Chapter 4. Indeed, given an approximate eigen-pair, we derive an explicit expression of the structured backward error under structured perturbation of $\mathrm{P} \in \mathbb{S}$. Note that, structured backward error is always greater than or equal to the unstructured backward error. We show that structured backward error is bounded above by a scalar multiple of the unstructured backward error for few approximate eigen-pairs. Besides, for each structure, there are certain approximate eigen-pairs for which they are equal.

Further using the expression of structured backward error of approximate eigen-pair of palindromic polynomial P, we obtain structured backward error of the corresponding approximate eigen-pair of its structured linearizations. Then we identify "good" structured linearization L of P that minimizes $\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L})/\eta(\lambda, x, \mathrm{P})$.

Its evident that structure preserving algorithms are required to produce symmetry in the computed eigenvalues. Structured preserving algorithms have been proposed in the literature to compute the eigenpair of a structured matrix polynomials, see [22, 44, 66]. We mention that the computable expressions of structured backward error obtained in this chapter has an important role to play in analyzing the backward stability of structured preserving algorithms.

Finally, we define structured backward error of an approximate eigenvalue of a structured matrix polynomial and apply it to establish a partial equality between structured and unstructured pseudospectra of H-palindromic / H-anti-palindromic polynomials.

The chapter is organized as follows. In section 5.2, we first discuss the eigensymmetry and backward error of an approximate eigenpair of palindromic polynomials. In section 5.3, we obtain expressions for structured backward error of approximate eigen-pair of palindromic polynomials. In section 5.4 we explain the palindromic linearizations of structure polynomials and determine "good" structured linearization.

5.2 Eigen-symmetry of palindromic polynomials

A matrix polynomial $\mathrm{P}(z) = \sum_{j=0}^{m} z^j A_j \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is called a $*$-palindromic or $*$-anti-palindromic if

$$\mathrm{P}^*(z) = z^m \mathrm{P}(1/z) \text{ or } \mathrm{P}^*(z) = z^m \mathrm{P}(-1/z),$$

respectively, for all $z \in \mathbb{C} \setminus \{0\}$, where $\mathrm{P}^*(z) = \sum_{j=0}^{m} z^j A_j^*$ and $* \in \{T, H\}$. It can be shown as in Theorem 3.1, that if λ is an eigenvalue of a $*$-palindromic polynomial P then $1/\lambda^*$ is also an eigenvalue of P. This $(\lambda, 1/\lambda^*)$ pairing is known as symplectic eigen-symmetry. Due to the structure of the coefficients of a palindromic polynomial, the spectrum of a $*$-palindromic polynomial inherits this spectral symmetry. Table 5.1 gives the eigen-symmetry of a palindromic polynomial.

TH-789_04612301

$\mathbb{S}$	**eigenvalue pair**	**eigentriple**
T-palindromic / T-antipalindromic	$(\lambda, 1/\lambda)$	$(\lambda, x, \overline{y}), (1/\lambda, y, \overline{x})$
H-palindromic/ H-anti-palindromic	$(\lambda, 1/\overline{\lambda})$	$(\lambda, x, y), (1/\overline{\lambda}, y, x)$

Table 5.1: Eigensymmetry of palindromic polynomials.

The validity of the Table 5.1 is followed by Table 3.1. We now show that, given $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and $\mathrm{P} \in \mathbb{S}$, there always exists a polynomial $\triangle\mathrm{P} \in \mathbb{S}$ such that $(\mathrm{P}(\lambda) + \triangle\mathrm{P}(\lambda))x = 0$, that is, (λ, x) is an eigen-pair of $\mathrm{P} + \triangle\mathrm{P}$. For $x \in \mathbb{C}^n$ with $\|x\|_2 = 1$, we define the projection $P_x := I - xx^H$.

Theorem 5.2.1. *Let $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ be a $*$-palindromic matrix polynomial, where $* \in \{T, H\}$. Suppose $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and set $r := -\mathrm{P}(\lambda)x$ and $\Lambda_m := [1, \lambda, \ldots, \lambda^m]^T$. Define*

$$\triangle A_j := \begin{cases} -\overline{x}x^T A_j x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^{m-j}}\,\overline{x}r^T P_x + \overline{\lambda^j} P_x^* r x^H], & \text{if } * = T, \\ -xx^H A_j x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m-j} x r^H P_x + \overline{\lambda^j} P_x r x^H], & \text{if } * = H. \end{cases}$$

$$\triangle A_{m-j} := \begin{cases} \triangle A_j^*,\ j = 0 : (m+1)/2, & \text{if } m \text{ is odd} \\ \triangle A_j^*,\ j = 0 : (m-2)/2, & \text{if } m \text{ is even.} \end{cases}$$

If m is even define

$$\triangle A_{m/2} := \begin{cases} -\overline{x}x^T A_{m/2} x x^H + \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}[\overline{x}r^T P_x + P_x^* r x^H], & \text{if } * = T, \\ -xx^H A_{m/2} x x^H + \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m/2} x r^H P_x + \overline{\lambda^{m/2}} P_x r x^H], & \text{if } * = H, \end{cases}$$

Then $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ where $\triangle\mathrm{P}$ is a $*$-palindromic polynomial.

Proof: The proof is computational and is easy to check.■

Next we consider $*$-anti-palindromic polynomials.

Theorem 5.2.2. *Let $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ be a $*$-anti-palindromic matrix polynomial, where $* \in \{T, H\}$. Suppose $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and set $r := -\mathrm{P}(\lambda)x$ and $\Lambda_m := [1, \lambda, \ldots, \lambda^m]^T$. Define*

$$\triangle A_j := \begin{cases} -\overline{x}x^T A_j x x^H - \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^{m-j}}\overline{x}r^T P_x - \overline{\lambda^j} P_x^* r x^H], & \text{if } * = T, \\ -xx^H A_j x x^H - \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m-j} x r^H P_x - \overline{\lambda^j} P_x r x^H], & \text{if } * = H. \end{cases}$$

$$\triangle A_{m-j} := \begin{cases} -\triangle A_j^*,\ j = 0 : (m+1)/2, & \text{if } m \text{ is odd,} \\ -\triangle A_j^*,\ j = 0 : (m-2)/2, & \text{if } m \text{ is even.} \end{cases}$$

TH-789_04612301

If m is even then define

$$\triangle A_{m/2} := \begin{cases} -\overline{x}x^T A_{m/2} x x^H - \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}[\overline{x}r^T P_x - P_x^* r x^H], & \text{if } * = T, \\ -x x^H A_{m/2} x x^H - \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m/2} x r^H P_x - \overline{\lambda^{m/2}} P_x r x^H], & \text{if } * = H. \end{cases}$$

Then $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\triangle\mathrm{P}$ *is a* $*$*-anti-palindromic polynomial.*

Proof: The proof is computational and is easy to check.■

5.3 Structured backward error of approximate eigenpair

In this section we derive structured backward error of an approximate eigenpair (λ, x) of $\mathrm{P} \in \mathbb{S}$. The backward error of (λ, x) is defined as the smallest, in norm, perturbation $\triangle\mathrm{P}$ of P such that (λ, x) is an eigenpair of $\mathrm{P} + \triangle\mathrm{P}$, that is, $(\mathrm{P}(\lambda) + \triangle\mathrm{P}(\lambda))x = 0$. We make the convention throughout the chapter that, $\triangle\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ is of the form $\triangle\mathrm{P} = \sum_{j=0}^m z^j \triangle A_j$. Recall that the Frobenius and the spectral norm defined on $\mathbb{P}_m(\mathbb{C}^{n\times n})$ are given by

$$\|\mathrm{P}\|_F := \left(\sum_{j=0}^m \|A_j\|_F^2\right)^{1/2} \quad \text{and } \|\mathrm{P}\|_2 := \left(\sum_{j=0}^m \|A_j\|_2^2\right)^{1/2}$$

respectively where $\mathrm{P}(z) = \sum_{i=0}^m z^j A_j$. Then it follows that, for any $\lambda \in \mathbb{C}, \|\mathrm{P}(\lambda)\| \leq \|\mathrm{P}\|_{F,2}\, \|\Lambda_m\|_2$, where $\Lambda_m = [1,\, \lambda,\, \ldots,\, \lambda^m]^T$.

By convention, if $(\lambda, x) \in \mathbb{C}\times\mathbb{C}^n$, then x is assumed to be nonzero, that is, $x \neq 0$. Treating (λ, x) as an approximate eigenpair of the polynomial $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$, we define the backward error of (λ, x) by

$$\begin{aligned} \eta_F(\lambda, x, \mathrm{P}) &:= \min_{\triangle\mathrm{P}\in\mathbb{P}_m(\mathbb{C}^{n\times n})} \{\|\triangle\mathrm{P}\|_F : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}, \\ \eta_2(\lambda, x, \mathrm{P}) &:= \min_{\triangle\mathrm{P}\in\mathbb{P}_m(\mathbb{C}^{n\times n})} \{\|\triangle\mathrm{P}\|_2 : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}. \end{aligned}$$

Setting $r := -\mathrm{P}(\lambda)x$, we have $\eta_F(\lambda, x, \mathrm{P}) = \|r\|_2/\|x\|_2\|\Lambda_m\|_2 = \eta_2(\lambda, x, \mathrm{P})$, see (4.1). Henceforth, we denote $\eta_F(\lambda, x, \mathrm{P})$ and $\eta_2(\lambda, x, \mathrm{P})$ by $\eta(\lambda, x, \mathrm{P})$.

Next assume that $\mathrm{P} \in \mathbb{S}$. Then we define the structured backward error of (λ, x) by

$$\begin{aligned} \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) &:= \min_{\triangle\mathrm{P}\in\mathbb{S}} \{\|\triangle\mathrm{P}\|_F : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\} \\ \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) &:= \min_{\triangle\mathrm{P}\in\mathbb{S}} \{\|\triangle\mathrm{P}\|_2 : \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0\}. \end{aligned}$$

Unless otherwise stated, we denote $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$ for both $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$ and $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

By Theorem 5.2.1 and Theorem 5.2.2 it is obvious to see that $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) < \infty$ and $\eta(\lambda, x, \mathrm{P}) \leq \eta^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

To make the presentation simple we define $\Pi_i : \mathbb{C}^m \to \mathbb{C}^m$, by

$$\Pi_i([x_0,\, x_1,\, x_1,\, \ldots,\, x_{m-1}]^T) := [x_0,\, x_1,\, x_2,\, \ldots,\, x_{i-1},\, 0,\, 0,\, \ldots,\, 0]^T.$$

TH-789_04612301

We consider flip operator $R = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{m \times m}$.

Theorem 5.3.1. *Let $\mathbb{S}$ be a space of T-palindromic polynomials and $\mathrm{P} \in \mathbb{S}$. Assume that (λ, x) is an approximate eigen-pair of P and set $r := -\mathrm{P}(\lambda)x$. Then we have the following*

1. *m is odd:*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\,\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = -1 \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq \pm 1 \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}) \text{ if } \lambda = \pm 1 \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}R\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}\Lambda_m\|^2\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1 \end{cases}$$

Let $E_j := \frac{\overline{\lambda^j} + \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2}(x^T r)\overline{x}x^H$, $F_j := \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}P_x^T r x^H + \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2}\overline{x}r^T P_x$. For $j = 0 : (m-1)/2$, define

$$\triangle A_j := \begin{cases} F_j, & \lambda = -1 \\ E_j + F_j, & \lambda \neq -1 \end{cases}$$

and $\triangle A_{m-j} = \triangle A_j^T$. Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is a unique polynomial such that $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Next, define

$$\triangle A_j := \begin{cases} F_j, & \text{if } \lambda = -1 \\ E_j + F_j - \frac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r}P_x^T r r^T P_x}{|\lambda^{m-j}|^2\ \|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\ E_j + F_j - \frac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r}P_x^T r r^T P_x}{|\lambda^j|^2\ \|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \lambda \neq -1 \end{cases}$$

for $j = 0 : (m-1)/2$ and $\triangle A_{m-j} = \triangle A_j^T$. Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is a polynomial such that $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

2. *m is even:*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{1}{\sqrt{m+1}}\sqrt{2\|r\|_2^2 - |x^T r|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = \pm 1 \\ \sqrt{\frac{(2\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2 - |\lambda^{m/2}|^2)|x^T r|^2}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4} + 2\frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq \pm 1 \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = \pm 1 \\ \sqrt{\frac{\|\Lambda_m + R\Lambda_m\|_2^2 - 3|\lambda^{m/2}|^2}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4}|x^T r|^2 + \frac{(2\|\Pi_{m/2}R\Lambda_m\|_2^2 + |\lambda^{m/2}|^2)\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1; \\ \sqrt{\frac{\|\Lambda_m + R\Lambda_m\|_2^2 - 3|\lambda^{m/2}|^2}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4}|x^T r|^2 + \frac{(2\|\Pi_{m/2}\Lambda_m\|_2^2 + |\lambda^{m/2}|^2)\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

TH-789_04612301

Set

$$
\begin{aligned}
G_j &:= \frac{\overline{\lambda^j} + \overline{\lambda^{m-j}}}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2}\overline{x}x^T r x^H + \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}P_x^T r x^H + \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2}\overline{x}r^T P_x \\
H_{m/2} &:= \frac{\overline{\lambda^{m/2}}}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2}\overline{x}x^T r x^H + \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}P_x^T r x^H + \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}\overline{x}r^T P_x.
\end{aligned}
$$

Now for $j = 0 : (m-2)/2$ *define*

$$
\begin{aligned}
\triangle A_j &:= \begin{cases} \frac{\lambda^j}{\|\Lambda_m\|_2^2}(x^T r)\overline{x}x^H + \frac{1}{\|\Lambda_m\|_2^2}[\lambda^j P_x^T r x^H + \lambda^{m-j}\overline{x}r^T P_x], & \text{if } \lambda = \pm 1 \\ G_j, & \text{if } \lambda \neq \pm 1 \end{cases} \\
\triangle A_{m/2} &:= \begin{cases} \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2}\,[(x^T r)\overline{x}x^H + P_x^T r x^H + \overline{x}r^T P_x], & \text{if } \lambda = \pm 1 \\ H_{m/2}, & \text{if } \lambda \neq \pm 1 \end{cases}
\end{aligned}
$$

and $\triangle A_{m-j} = \triangle A_j^T$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ *is a unique polynomial such that* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Further, for $j = 0 : (m-2)/2$ *define*

$$
\triangle A_j := \begin{cases} \frac{\lambda^j}{\|\Lambda_m\|_2^2}(x^T r)\overline{x}x^H + \frac{1}{\|\Lambda_m\|_2^2}[\lambda^j P_x^T r x^H + \lambda^{m-j}\overline{x}r^T P_x] - \frac{\lambda^{m-j}\ \overline{x^T r}P_x^T r r^T P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } \lambda = \pm 1 \\[2ex] G_j - \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r}\ P_x^T r r^T P_x}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2\ |\lambda^{m-j}|^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\[2ex] G_j - \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r}\ P_x^T r r^T P_x}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2\ |\lambda^j|^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \lambda \neq \pm 1 \end{cases}
$$

$$
\triangle A_{m/2} := \begin{cases} \dfrac{\lambda^{m/2}}{\|\Lambda_m\|_2^2}\,[(x^T r)\overline{x}x^H + P_x^T r x^H + \overline{x}r^T P_x] - \dfrac{\lambda^{m/2}\overline{x^T r}\ P_x^T r r^T P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } \lambda = \pm 1 \\[2ex] H_{m/2} - \dfrac{\overline{\lambda^{m/2}}\ \overline{x^T r}\ P_x^T r r^T P_x}{\|\Pi_{(m/2)+1}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } \lambda \neq \pm 1, \end{cases}
$$

and $\triangle A_{m-j} = \triangle A_j^T$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^{m} z^j \triangle A_j$ *is a polynomial such that a* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: First suppose that m is odd. By Theorem 5.2.1 there exists a polynomial $\triangle \mathrm{P} \in \mathbb{S}$ such that $\triangle\mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. For $j = 0 : (m-1)/2$, consider

$$
\widetilde{\triangle A_j} := Q^T \triangle A Q = \begin{pmatrix} a_{jj} & a_j^T \\ b_j & X_j \end{pmatrix} \text{ and } \triangle A_j^T = \triangle A_{m-j},
$$

where $Q := [x, Q_1]$ is a unitary matrix. Now since $\triangle\mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have,

$$
\begin{pmatrix} \sum_{j=0}^{m} \lambda^j a_{jj} \\ \sum_{j=0}^{(m-1)/2} \lambda^j b_j + \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.
$$

The minimum norm solution of $\sum_{j=0}^{(m-1)/2} \lambda^j b_j + \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j = Q_1^T r$ is given by $b_j =$

TH-789_04612301

$\frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r$ and $a_j = \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^T r.$

For $\lambda = -1$, we have $x^T r = 0$. Hence the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^T r$, is $a_{jj} = 0$. So we have $\triangle A_j$ for $j = 0 : m$, as follows:

$$\triangle A_j = \overline{Q} \begin{pmatrix} 0 & (\frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^T r)^T \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r & X_j \end{pmatrix} Q^H, \quad \triangle A_j^T = \triangle A_{m-j}, \; j = 0 : (m-1)/2.$$

Setting $X_j = 0$, $j = 0 : m$, we obtain $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\, \frac{\|r\|_2}{\sqrt{m+1}} = \sqrt{2}\, \eta(\lambda, x, \mathrm{P})$. Simplifying the expressions of $\triangle A_j$, $j = 0 : m$, we obtain the desired result.

Now let $\lambda \neq -1$. Notice that $a_{jj} = a_{m-j,m-j}$, $j = 0 : (m-1)/2$. Hence the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^T r$ which implies $\sum_{j=0}^{(m-1)/2} (\lambda^j + \lambda^{m-j}) a_{jj} = x^T r$, is

$$a_{jj} = \frac{\overline{\lambda^j} + \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} x^T r.$$

Thus we obtain

$$\begin{aligned} \triangle A_j &= \overline{Q} \begin{pmatrix} \frac{\overline{\lambda^j} + \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} x^T r & (\frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^T r)^T \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r & X_j \end{pmatrix} Q^H, \quad j = 0 : (m-1)/2 \\ \triangle A_{m-j} &= A_j^T, j = 0 : (m-1)/2. \end{aligned}$$

Now setting $X_j = 0, j = 0 : (m-1)/2$, we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2} \sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expression of $\triangle A_j$, $i = 0 : m$ we obtain the desired result.

In particular, when $\lambda = 1$, we get $\|\Lambda_m\|_2^2 = m+1$ and $\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2 = 2(m+1)$, which gives,

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |x^T r|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}).$$

Now we consider spectral norm. For $\lambda = -1$, by DKW Theorem 1.2.5 applied to $\triangle A_j$ gives $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2^2}$ and $X_j = 0, j = 0 : (m-1)/2$. Hence $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\sqrt{m+1}} = \eta(\lambda, x, \mathrm{P})$. Thus we obtain the same $\triangle A_j$ that we have obtained for Frobenius norm.

Next let $\lambda \neq -1$. For $j = 0 : (m-1)/2$, applying DKW Theorem 1.2.5 to $\triangle A_j$, we have

$$\mu_{\triangle A_j} = \begin{cases} \sqrt{\frac{|\lambda^j + \lambda^{m-j}|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^4} |x^T r|^2 + \frac{|\lambda^{m-j}|^2 \, (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{\frac{|\lambda^j + \lambda^{m-j}|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^4} |x^T r|^2 + \frac{|\lambda^j|^2 \, (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1, \end{cases}$$

TH-789_04612301

and

$$X_j = \begin{cases} -\dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r} Q_1^T r (Q_1^T r)^T}{|\lambda^{m-j}|^2\ \|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\ -\dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} + |\lambda^{m-j}|^2\overline{\lambda^j})\ \overline{x^T r} Q_1^T r (Q_1^T r)^T}{|\lambda^j|^2\ \|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1. \end{cases}$$

This gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}\, R\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}\Lambda_m\|^2\ (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

Simplifying the expression of $\triangle A_j, \quad j = 0 : (m-1)/2$ we obtain the desired result.

Now if $\lambda = 1$, we have $\|\Lambda_m\|_2^2 = m+1$ and $\|\Pi_{(m+1)/2}(\Lambda_m + R\Lambda_m)\|_2^2 = 2(m+1)$, which gives,

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) \quad = \quad \frac{\|r\|_2}{\sqrt{m+1}} = \eta(\lambda, x, \mathrm{P}).$$

Next, suppose that m is even. Note that, $A_{m/2} = A_{m/2}^T$. Then we have

$$\begin{pmatrix} \sum_{j=0}^{m} \lambda^j a_{jj} \\ \sum_{j=0}^{(m-2)/2} \lambda^j b_j + \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

The minimum norm solution of $\sum_{j=0}^{(m-2)/2} \lambda^j b_j + \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} = Q_1^T r$ is given by

$$b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r,\ a_j = \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^T r,\ a_{m/2} = \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} Q_1^T r.$$

For $\lambda = \pm 1$, the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^T r$ is $a_{jj} = \frac{\lambda^j}{\|\Lambda_m\|_2^2} x^T r$. Hence we have,

$$\triangle A_j = \overline{Q} \begin{pmatrix} \frac{\lambda^j}{\|\Lambda_m\|_2^2} x^T r & \frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2} (Q_1^T r)^T \\ \frac{\lambda^j}{\|\Lambda_m\|_2^2} Q_1^T r & X_j \end{pmatrix} Q^H,\ \triangle A_{m/2} = \overline{Q} \begin{pmatrix} \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2} x^T r & \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2} (Q_1^T r)^T \\ \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2} Q_1^T r & X_{m/2} \end{pmatrix} Q^H,$$

and $\triangle A_{m-j} = \triangle A_j^T,\ j = 0 : (m-2)/2$. Setting $X_j = 0,\ j = 0 : m$, we have $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |x^T r|^2}$. Simplifying $\triangle A_j$s for $j = 0 : m$ we obtain the desired result.

Again, if $\lambda \neq \pm 1$, the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^T r$ is

$$a_{jj} = \frac{\overline{\lambda^j} + \overline{\lambda^{m-j}}}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2} R\Lambda_m\|_2^2} x^T r,\ a_{m/2,m/2} = \frac{\overline{\lambda^{m/2}}}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2} R\Lambda_m\|_2^2} x^T r.$$

TH-789_04612301

Hence we have

$$\begin{aligned}
\triangle A_j &= \overline{Q} \begin{pmatrix} \frac{\overline{\lambda^j}+\overline{\lambda^{m-j}}}{\|\Pi_{(m+2)/2}\Lambda_m+\Pi_{m/2}R\Lambda_m\|_2^2} x^T r & \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2}(Q_1^T r)^T \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r & X_j \end{pmatrix} Q^H, \; j = 0 : (m-2)/2, \\
\triangle A_{m/2} &= \overline{Q} \begin{pmatrix} \frac{\overline{\lambda^{m/2}}}{\|\Pi_{(m+2)/2}\Lambda_m+\Pi_{m/2}R\Lambda_m\|_2^2} x^T r & \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}(Q_1^T r)^T \\ \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} Q_1^T r & X_{m/2} \end{pmatrix} Q^H \\
\triangle A_{m-j} &= \triangle A_j^T, \;\; j = 0 : (m-2)/2.
\end{aligned}$$

Setting $X_j = 0, \; j = 0 : m$, we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{\frac{(2\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2 - |\lambda^{m/2}|^2)|x^T r|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4} + 2\frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}.$$

which gives the desired result. Simplifying the expressions of $\triangle A_j$ the desired result follows.

Next we consider spectral norm. If $\lambda = \pm 1$, applying DKW Theorem 1.2.5 to $\triangle A_j$ we have, $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2^2}$ for $j = 0 : m$ and $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\sqrt{m+1}} = \eta(\lambda, x, \mathrm{P})$. Now by DKW Theorem 1.2.5, we have

$$X_j = -\frac{\lambda^{m-j} \; Q_1^T r (Q_1^T r)^T}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^T r|^2)}, \; j = 0 : (m-2)/2, \; X_{m/2} = -\frac{\lambda^{m/2} \; Q_1^T r (Q_1^T r)^T}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^T r|^2)}.$$

Therefore, we have

$$\begin{aligned}
\triangle A_j &= \frac{\lambda^j}{\|\Lambda_m\|_2^2} r x^H + \frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2} \overline{x} r^T P_x - \frac{\lambda^{m-j} \; P_x^T r r^T P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^T r|^2)}, \\
\triangle A_{m/2} &= \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2} \left[r x^H + \overline{x} r^T P_x \right] - \frac{\lambda^{m/2} \; P_x^T r r^T P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^T r|^2)}.
\end{aligned}$$

Again for $\lambda \neq \pm 1$, we have

$$\begin{aligned}
\mu_{\triangle A_j} &= \begin{cases} \sqrt{\frac{|\lambda^j + \lambda^{m-j}|^2 \; |x^T r|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4} + \frac{|\lambda^{m-j}|^2 \; (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{\frac{|\lambda^j + \lambda^{m-j}|^2 \; |x^T r|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4} + \frac{|\lambda^j|^2 \; (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1 \end{cases} \\
\mu_{\triangle A_{m/2}} &= \sqrt{\frac{|\lambda^{m/2}|^2 \; |x^T r|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4} + \frac{|\lambda^{m/2}|^2 \; (\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}.
\end{aligned}$$

Consequently by DKW Theorem 1.2.5 we have

$$\begin{aligned}
X_j &= \begin{cases} -\frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} + |\lambda^{m-j}|^2 \overline{\lambda}^j) \; \overline{x^T r} \; Q_1^T r (Q_1^T r)^T}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2 \; |\lambda^{m-j}|^2 \; (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\ -\frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} + |\lambda^{2m-j}|^2 \overline{\lambda^j}) \; \overline{x^T r} \; Q_1^T r (Q_1^T r)^T}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2 \; |\lambda^j|^2 \; (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1; \end{cases} \\
X_{m/2} &= -\frac{\overline{\lambda^{m/2}} \; \overline{x^T r} \; Q_1^T r (Q_1^T r)^T}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^T r|^2)}.
\end{aligned}$$

TH-789_04612301

Hence we obtain

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{\frac{\|\Lambda_m + R\Lambda_m\|_2^2 - 3|\lambda^{m/2}|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4}|x^Tr|^2 + \frac{2\|\Pi_{m/2}R\Lambda_m\|_2^2 + |\lambda^{m/2}|^2\ (\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}, \\ \hfill \text{if } |\lambda| > 1 \\ \sqrt{\frac{\|\Lambda_m + R\Lambda_m\|_2^2 - 3|\lambda^{m/2}|^2}{\|\Pi_{(m+2)/2}\Lambda_m + \Pi_{m/2}R\Lambda_m\|_2^4}|x^Tr|^2 + \frac{2\|\Pi_{m/2}\Lambda_m\|_2^2 + |\lambda^{m/2}|^2\ (\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}, \\ \hfill \text{if } |\lambda| \leq 1. \end{cases}$$

Simplifying the expressions of $\triangle A_j$s we obtain the desired result.■

Note that if $Y \in \mathbb{C}^{n\times n}$ is such that $Yx = 0$ and $Y^Tx = 0$ then $Y = (I - xx^H)^T Z(I - xx^H)$ for some matrix Z. Hence from the proof of Theorem 5.3.1, we obtain that if K is a T-palindromic polynomial such that $\mathrm{P}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K}(z) = \triangle\mathrm{P}(z) + (I - xx^H)^T\mathrm{N}(z)(I - xx^H)$ for some T-palindromic polynomial N, where $\triangle\mathrm{P}$ is given in Theorem 5.3.1.

Remark 5.3.2. *If $|x^Tr| = \|r\|_2$, then $\|Q_1^Tr\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

Theorem 5.3.3. *Let $\mathbb{S}$ be the space of T-anti-palindromic polynomials and $\mathrm{P} \in \mathbb{S}$. Assume that $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ and set $r := -\mathrm{P}(\lambda)x$. Then we have the following.*

1. m is odd: we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\ \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = 1 \\ \sqrt{2}\sqrt{\frac{|x^Tr|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^Tr|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta_2(\lambda, x, \mathrm{P}), & \text{if } \lambda = 1 \\ \sqrt{2}\sqrt{\frac{|x^Tr|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}R\Lambda_m\|_2^2(\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}, \\ \hfill \text{if } |\lambda| > 1, \\ \sqrt{2}\sqrt{\frac{|x^Tr|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}\Lambda_m\|_2^2(\|r\|_2^2 - |x^Tr|^2)}{\|\Lambda_m\|_2^4}}, \\ \hfill \text{if } |\lambda| \leq 1. \end{cases}$$

Set

$$E_j := \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2}(x^Tr)\overline{x}x^H - \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^{m-j}}\overline{x}r^TP_x - \overline{\lambda^j}P_x^Trx^H]$$

$$F_j := \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2}(x^Tr)\overline{x}x^H - \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^{m-j}}\overline{x}r^TP_x - \overline{\lambda}^jP_x^Trx^H].$$

For $j = 0 : (m-1)/2$, define

$$\triangle A_j := \begin{cases} \frac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ E_j, & \lambda \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -A_j^T$. Then $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is a unique polynomial $\triangle\mathrm{P} \in \mathbb{S}$ such

TH-789_04612301

that $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. *Next, define*

$$\triangle A_j := \begin{cases} \dfrac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ E_j + \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} - |\lambda^{m-j}|^2\overline{\lambda^j})\overline{x^T r}P_x^T rr^T P_x}{|\lambda^{m-j}|^2\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\ E_j + \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} - |\lambda^{m-j}|^2\overline{\lambda^j})\overline{x^T r}P_x^T rr^T P_x}{|\lambda^j|^2\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \lambda \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -A_j^T$, $j = 0 : (m-1)/2$. *Then* $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a polynomial such that* $\triangle\mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

2. *m is even:*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\;\eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = 1 \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } \lambda \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } \lambda = 1, \\ \sqrt{\frac{2|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{(2\|\Pi_{m/2}R\Lambda_m\|^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1, \\ \sqrt{\frac{2|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{(2\|\Pi_{m/2}\Lambda_m\|^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

For $j = 0 : (m-2)/2$, *consider*

$$\triangle A_j := \begin{cases} \frac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ F_j, & \text{if } \lambda \neq 1 \end{cases}, \quad \triangle A_{m/2} := \begin{cases} \frac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -A_j^T$, $j = 0 : (m-2)/2$. *Then* $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a unique polynomial such that* $\triangle\mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. *Next, define*

$$\triangle A_j := \begin{cases} \dfrac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ F_j + \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} - |\lambda^{m-j}|^2\overline{\lambda^j})\overline{x^T r}P_x^T rr^T P_x}{|\lambda^{m-j}|^2\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1 \\ F_j + \dfrac{(|\lambda^j|^2\overline{\lambda^{m-j}} - |\lambda^{m-j}|^2\overline{\lambda^j})\overline{x^T r}P_x^T rr^T P_x}{|\lambda^j|^2\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \lambda \neq 1 \end{cases}$$

$$\triangle A_{m/2} := \begin{cases} \dfrac{1}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda = 1 \\ \dfrac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}[rx^H - \overline{x}r^T], & \text{if } \lambda \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -A_j^T$, $j = 0 : (m-2)/2$. *Then* $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a polynomial such that* $\triangle\mathrm{P} \in \mathbb{S}, \triangle\mathrm{P} \in \mathbb{S}$ *such that* $\mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

TH-789_04612301

Proof: First assume that $\mathrm{P} \in \mathbb{S}$ and m is odd. By Theorem 5.2.2 there exists $\triangle\mathrm{P} \in \mathbb{S}$ such that $\triangle\mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. For $j = 0 : (m-1)/2$, consider

$$\widetilde{\triangle A_j} := Q^T \triangle A Q = \begin{pmatrix} a_{jj} & a_j^T \\ b_j & X_j \end{pmatrix} \text{ and } \triangle A_{m-j} = -\triangle A_j^T, \; j = 0 : (m-1)/2,$$

where $Q = [x, \, Q_1]$ is a unitary matrix. Since $\triangle\mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have,

$$\begin{pmatrix} \sum_{j=0}^{(m-1)/2} \lambda^j a_{jj} - \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_{jj} \\ \\ \sum_{j=0}^{(m-1)/2} \lambda^j b_j - \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

The minimum norm solution of $\sum_{j=0}^{(m-1)/2} \lambda^j b_j - \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j = Q_1^T r$ is given by $b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^T r$ and $a_j = -\frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^T r$, $j = 0 : (m-1)/2$.

For $\lambda = 1$, we have $x^T r = 0$. Hence the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^T r$, is $a_{jj} = 0$, $j = 0 : m$. Thus we have

$$\triangle A_j = \overline{Q} \begin{pmatrix} 0 & -\frac{\overline{\lambda^{m-j}} (Q_1^T r)^T}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \; A_{m-j} = A_j^T \; j = 0 : (m-1)/2.$$

Setting $X_j = 0, j = 0 : m$ we have the minimum Frobenius norm of $\triangle A_j$. Consequently we have,

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\, \frac{\|r\|_2}{\|\Lambda_m\|_2} = \sqrt{2}\, \eta(\lambda, x, \mathrm{P}).$$

Simplifying the expression of $\triangle A_j, j = 0 : m$ we obtain the desired result.

Next suppose that $\lambda \neq 1$. Then the minimum norm solution of

$$\sum_{j=0}^{(m-1)/2} \lambda^j a_{jj} - \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_{jj} = x^T r$$

is given by

$$a_{jj} = \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} x^T r, \; j = 0 : (m-1)/2.$$

Therefore we have

$$\triangle A_j = \overline{Q} \begin{pmatrix} \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} x^T r & -\frac{\overline{\lambda^{m-j}} (Q_1^T r)^T}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \; j = 0 : (m-1)/2.$$

Taking $X_j = 0$, we obtain the minimum Frobenius norm of $\triangle A_j$ and consequently we have

$$\begin{aligned} \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) &= \sqrt{2} \sqrt{\frac{\sum_{j=0}^{(m-1)/2} |\lambda^j - \lambda^{m-j}|^2 |x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^4} + \frac{\sum_{j=0}^{(m-1)/2} (|\lambda^{m-j}|^2 + |\lambda^j|^2)(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}} \\ &= \sqrt{2} \sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}. \end{aligned}$$

TH-789_04612301

Simplifying the expression of $\triangle A_j$ for all $j = 0 : (m-1)/2$ we obtain the desired result.

Next we consider the spectral norm. Assume that $\lambda = 1$. Applying DKW Theorem 1.2.5 to $\triangle A_j$ we have $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2}$, $j = 0 : (m-1)/2$ and $X_j = 0$. Therefore we have $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|\Lambda_m\|_2} = \eta(\lambda, x, \mathrm{P})$. Simplifying the expression of $\triangle A_j$ we obtain $\triangle A_j = -\frac{1}{\|\Lambda_m\|_2^2}[\overline{x}r^T - rx^H]$. Hence the result follows.

Next, suppose that $\lambda \neq 1$. In this case we have

$$\mu_{\triangle A_j} = \begin{cases} \sqrt{\frac{|\lambda^j - \lambda^{m-j}|^2 |x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^4} + \frac{|\lambda^{m-j}|^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if} \quad |\lambda| > 1 \\ \sqrt{\frac{|\lambda^j - \lambda^{m-j}|^2 |x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^4} + \frac{|\lambda^{j}|^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if} \quad |\lambda| \leq 1 \end{cases}$$

for $j = 0 : (m-1)/2$. Therefore applying DKW Theorem 1.2.5 to $\triangle A_j$ we have

$$X_j = \begin{cases} \frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} - |\lambda^{m-j}|^2 \overline{\lambda^j})\overline{x^T r}(Q_1^T r)(Q_1^T r)^T}{|\lambda^{m-j}|^2 \|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1, \\ \frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} - |\lambda^{m-j}|^2 \overline{\lambda^j})\overline{x^T r}(Q_1^T r)(Q_1^T r)^T}{|\lambda^{j}|^2 \|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2(\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \end{cases}$$

for $j = 0 : (m-1)/2$.

Hence we have

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}R\Lambda_m\|_2^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if} \quad |\lambda| > 1, \\ \sqrt{2}\sqrt{\frac{|x^T r|^2}{\|\Pi_{(m+1)/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|\Pi_{(m+1)/2}\Lambda_m\|_2^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if} \quad |\lambda| \leq 1. \end{cases}$$

Simplifying the expression of $\triangle A_j$ for $j = 0 : (m-1)/2$ we obtain the desired result.

Next, assume that m is even. Note that in this case $A_{m/2}^T = -A_{m/2}$. Consequently we have

$$\begin{pmatrix} \sum_{j=0}^{(m-2)/2} \lambda^j a_{jj} - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_{jj} \\ \sum_{j=0}^{(m-2)/2} \lambda^j b_j + a_{m/2}\lambda^{m/2} - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j \end{pmatrix} = \begin{pmatrix} x^T r \\ Q_1^T r \end{pmatrix}.$$

Note that $a_{m/2,m/2} = 0$, since $A_{m/2}$ is a skew-symmetric matrix. The minimum norm solution of $\sum_{j=0}^{(m-2)/2} \lambda^j b_j + a_{m/2}\lambda^{m/2} - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j = Q_1^T r$ is given by

$$b_j = \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2}, \; a_j = -\frac{\overline{\lambda^{m-j}} Q_1^T r}{\|\Lambda_m\|_2^2}, \; a_{m/2} = \frac{\overline{\lambda^{m/2}} Q_1^T r}{\|\Lambda_m\|_2^2}, j = 0 : (m-2)/2.$$

For $\lambda = 1$, $x^T r = 0$. Hence the minimum norm solution of $\sum_{j=0}^{(m-2)/2} \lambda^j a_{jj} - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_{jj} = 0$ is given by $a_{jj} = 0$. Consequently, we have,

$$\triangle A_j = \overline{Q} \begin{pmatrix} 0 & -\frac{(Q_1^T r)^T}{\|\Lambda_m\|_2^2} \\ \frac{Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \; j = 0 : (m-2)/2.$$

Setting $X_j = 0, j = 0 : (m-2)/2$ we obtain the minimum norm of $\triangle A_j$ and consequently we have $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2}\,\eta(\lambda, x, \mathrm{P})$. Simplifying the expression of $\triangle A_j$ we obtain the desired

TH-789_04612301

result.

Next, suppose that $\lambda \neq 1$. Then the minimum norm solution of

$$\sum_{j=0}^{(m-2)/2} \lambda^j a_{jj} - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_{jj} = x^T r$$

is given by

$$a_{jj} = \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} x^T r$$

and $a_{m/2,m/2} = 0$. Consequently, for $j = 0 : (m-2)/2$, we obtain

$$\begin{aligned} \triangle A_j &= \overline{Q} \begin{pmatrix} \frac{\overline{\lambda^j} - \overline{\lambda^{m-j}}}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} x^T r & -\frac{\overline{\lambda^{m-j}}(Q_1^T r)^T}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^j} Q_1^T r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \\ \triangle A_{m/2} &= \overline{Q} \begin{pmatrix} 0 & -\frac{\overline{\lambda^{m/2}}(Q_1^T r)^T}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^{m/2}} Q_1^T r}{\|\Lambda_m\|_2^2} & X_{m/2} \end{pmatrix} Q^H. \end{aligned}$$

Setting $X_j = 0 = X_{m/2}$, we get the minimum Frobenius norm of $\triangle A_j$ and $\triangle A_{m/2}$ respectively and these give,

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2} \sqrt{\frac{|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{\|r\|_2^2 - |x^T r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expressions of $\triangle A_j$ for $j = 0 : m/2$ we obtain the desired result.

Next consider spectral norm. Suppose that $\lambda = 1$. Then consider $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2}$ and by DKW Theorem 1.2.5, we have $X_j = 0$. Therefore we have $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\|\Lambda_m\|_2} = \eta(\lambda, x, \mathrm{P})$. Simplifying the expression of $\triangle A_j$ we obtain $\triangle A_j = -\frac{1}{\|\Lambda_m\|_2^2}[\overline{x} r^T - r x^H]$ for $j = 0 : m/2$. Hence the result follows.

Next, suppose $\lambda \neq 1$. Consider $\mu_{\triangle A_{m/2}} = \sqrt{\frac{|\lambda^m|^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}$. Then again by DKW Theorem 1.2.5 we have $X_{m/2} = 0$. Now for $j = 0 : (m-2)/2$ consider

$$\mu_{\triangle A_j} = \begin{cases} \sqrt{\frac{|\lambda^j - \lambda^{m-j}|^2 |x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^4} + \frac{|\lambda^{m-j}|^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1, \\ \sqrt{\frac{|\lambda^j - \lambda^{m-j}|^2 |x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^4} + \frac{|\lambda^j|^2(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

Then by DKW Theorem 1.2.5 applied to $\triangle A_j$ we have

$$X_j = \begin{cases} \frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} - |\lambda^{m-j}|^2 \overline{\lambda^j}) \overline{x^T r} (Q_1^T r)(Q_1^T r)^T}{|\lambda^{m-j}|^2 \|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2 (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| > 1, \\ \frac{(|\lambda^j|^2 \overline{\lambda^{m-j}} - |\lambda^{m-j}|^2 \overline{\lambda^j}) \overline{x^T r} (Q_1^T r)(Q_1^T r)^T}{|\lambda^j|^2 \|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2 (\|r\|_2^2 - |x^T r|^2)}, & \text{if } |\lambda| \leq 1, \end{cases}$$

TH-789_04612301

for $j = 0 : (m-2)/2$. Therefore we have

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{\frac{2|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{(2\|\Pi_{m/2} R\Lambda_m\|^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1, \\ \sqrt{\frac{2|x^T r|^2}{\|\Pi_{m/2}(\Lambda_m - R\Lambda_m)\|_2^2} + \frac{(2\|\Pi_{m/2} \Lambda_m\|^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^T r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| \leq 1. \end{cases}$$

Simplifying the expressions of $\triangle A_j$ for $j = 0 : m/2$ we obtain the desired result. ■

Analogous to the T-palindromic polynomial we conclude that if K is a T-anti-palindromic polynomial such that $\mathrm{P}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K}(z) = \triangle\mathrm{P}(z) + (I - xx^H)^T \mathrm{N}(z)(I - xx^H)$ for some T-anti-palindromic polynomial N, where $\triangle\mathrm{P}$ is given in Theorem 5.3.3.

Remark 5.3.4. *If $|x^T r| = \|r\|_2$, then $\|Q_1^T r\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

Next we consider H-palindromic/ H-anti-palindromic polynomials. To make the presentation simple we proceed as follows.

Let $z = a + i\,b \in \mathbb{C}$. Define a map vec : $\mathbb{C} \to \mathbb{R}^2$ by $\mathrm{vec}(z) = \begin{pmatrix} \mathsf{re}(z) \\ \mathsf{im}(z) \end{pmatrix}$. Further define a map $\mathtt{M} : \mathbb{C} \to \mathbb{R}^{2\times 2}$ by $\mathtt{M}(z) = \begin{pmatrix} \mathsf{re}(z) & -\mathsf{im}(z) \\ \mathsf{im}(z) & \mathsf{re}(z) \end{pmatrix}$. Then the following hold:

- $\mathrm{vec}(\overline{z}) = \Sigma\, \mathrm{vec}(z)$, where $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
- $\mathrm{vec}(z_1 z_2) = \mathtt{M}(z_1)\mathrm{vec}(z_2)$, $z_1, z_2 \in \mathbb{C}$.
- $\mathtt{M}(\overline{z}) = \mathtt{M}(z)^T$.

Assume that $\lambda \in \mathbb{C}$, and $a_{jj} \in \mathbb{C}$. Then if we have $\sum_{j=0}^m \lambda^i a_{jj} = x^H r$ then applying the map vec both sides we obtain

$$\mathrm{vec}(\sum_{j=0}^m \lambda^i a_{jj}) = \mathrm{vec}(x^H r) \;\Rightarrow\; \sum_{j=0}^m \mathtt{M}(\lambda^i)\mathrm{vec}(a_{jj}) = \mathrm{vec}(x^H r). \tag{5.1}$$

Theorem 5.3.5. *Let $\mathbb{S}$ be the space of H-palindromic polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Setting $r := -\mathrm{P}(\lambda)x$, we have the following.*

1. *Let m be odd. Then*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } |\lambda| \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|\Pi_{(m+1)/2} R\Lambda_m\|^2 \; (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|\Pi_{(m+1)/2} \Lambda_m\|^2 \; (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

TH-789_04612301

where $\widehat{r} = \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-1)/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r)$ *and*

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) \\ \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}, j = 0 : (m-1)/2.$$

Let $E_j := \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m-j} x r^H P_x + \overline{\lambda^j} P_x r x^H]$. *For* $j = 0 : (m-1)/2$ *define*

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + E_j, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + E_j, & \text{if } |\lambda| \neq 1 \end{cases}$$

and $\triangle A_{m-j} = \triangle A_j^H$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a unique polynomial such that* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. *Next define*

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + E_j - \frac{x^H r \; \lambda^{m-j} P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + E_j - \frac{\overline{e_j^T \widehat{r}} \; \overline{\lambda^j} \; \lambda^{m-j} \; P_x r r^H P_x}{|\lambda^{m-j}|^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ e_j^T \widehat{r} x x^H + E_j - \frac{\overline{e_j^T \widehat{r}} \; \overline{\lambda^j} \; \lambda^{m-j} \; P_x r r^H P_x}{|\lambda^i|^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1 \end{cases}$$

for $j = 0 : (m-1)/2$ *with* $\triangle A_{m-j} = \triangle A_j^H$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a polynomial such that* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

2. *Let* m *be even. Then*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + 2\frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } |\lambda| \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}) & \text{if } |\lambda| = 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} R \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

where $\widehat{r} = \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-2)/2} & H_{m/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r)$ *and*

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j} \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}, j = 0 : (m-2)/2,$$

$$H_{m/2} = \begin{pmatrix} \mathsf{re}\,(\lambda^{m/2}) \\ \mathsf{im}\,(\lambda^{m/2}) \end{pmatrix}.$$

Let $F_j := \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m-j} x r^H P_x + \overline{\lambda^j} P_x r x^H]$ *and* $G_{m/2} := \frac{1}{\|\Lambda_m\|_2^2}[\lambda^{m/2} x r^H P_x + \overline{\lambda^{m/2}} P_x r x^H]$.

TH-789_04612301

Then for $j = 0 : (m-2)/2$, define

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + F_j, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + F_j, & \text{if } |\lambda| \neq 1 \end{cases}, \quad \triangle A_{m/2} := \begin{cases} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} x r^H x x^H + G_{m/2}, \\ \qquad\qquad \text{if } |\lambda| = 1 \\ e_{m/2}^T \widehat{r} x x^H + G_{m/2}, \\ \qquad\qquad \text{if } |\lambda| \neq 1 \end{cases}$$

and $\triangle A_{m-j} = \triangle A_j^H$. Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is a unique polynomial such that $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. Next define

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x x^H r x^H + F_j - \frac{\lambda^{m-j} \; r^H x \; P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |r^H x|^2)}, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + F_j - \frac{\overline{e_j^T \widehat{r}} \; \overline{\lambda^j} \; \lambda^{m-j} \; P_x r r^H P_x}{|\lambda^{m-j}|^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ e_j^T \widehat{r} x x^H + F_j - \frac{\overline{e_j^T \widehat{r}} \; \overline{\lambda^j} \; \lambda^{m-j} \; P_x r r^H P_x}{|\lambda^i|^2 \; (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1 \end{cases}$$

$$\triangle A_{m/2} := \begin{cases} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} x x^H r x^H + G_m - \frac{\lambda^{m/2} \; r^H x \; P_x r r^H P_x}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |r^H x|^2)}, & \text{if } |\lambda| = 1 \\ e_{m/2}^T \widehat{r} x x^H + G_{m/2} - \frac{\overline{e_{m/2}^T \widehat{r}} \; P_x r r^H P_x}{(\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| \neq 1. \end{cases}$$

and $\triangle A_{m-j} = \triangle A_j^H$. Then $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ is a polynomial such that $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ and $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: Let $\mathrm{P} \in \mathbb{S}$ given by $\mathrm{P}(z) = \sum_{j=0}^m z^j A_j$. First suppose that m is odd. By Theorem 5.2.2 there exists $\triangle \mathrm{P} \in \mathbb{S}$, such that $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. Define

$$\triangle A_j := Q \begin{pmatrix} a_{jj} & a_j^H \\ b_j & X_j \end{pmatrix} Q^H, \quad X_{m-j} = X_j^H, \; j = 0 : (m-1)/2,$$

where $Q = [x, \, Q_1]$ is a unitary matrix. Since $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have

$$\begin{pmatrix} \sum_{j=0}^m \lambda^j a_{jj} \\ \sum_{j=0}^{(m-1)/2} \lambda^j b_j + \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}.$$

Consequently, the minimum norm solution of $\sum_{j=0}^{(m-1)/2} \lambda^j b_j + \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j = Q_1^H r$ is given by

$$b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r, \; a_j = \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^H r, j = 0 : (m-1)/2.$$

Note that for $|\lambda| = 1$, we have $\overline{x^H r} = \overline{\lambda^m} x^H r$. Hence the minimum norm solution of

TH-789_04612301

$\sum_{j=0}^{m} \lambda^j a_{jj} = x^H r$ is $a_{jj} = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r,\ j = 0 : m$. Therefore we have

$$\begin{aligned} \triangle A_j &= Q \begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r & \frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2} (Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H, \\ \triangle A_{m-j} &= \triangle A_j^H,\ j = 0 : (m-1)/2. \end{aligned}$$

Setting $X_j = 0$ gives $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2}$. Simplifying the expressions for $\triangle A_j$s we obtain the desired result.

Next, let $|\lambda| \neq 1$. The value of a_{jj} for the equation $\sum_{j=0}^{m} \lambda^j a_{jj} = x^H r,\ j = 0 : (m-1)/2$ is achieved after employing the condition $\bar{a}_{jj} = a_{m-j,m-j}$ $j = 0 : (m-1)/2$ in equation (5.1) which becomes

$$\sum_{j=0}^{(m-1)/2} (\mathtt{M}(\lambda^j) + \mathtt{M}(\lambda^{m-j})\Sigma)\mathrm{vec}(a_{jj}) = \mathrm{vec}(x^H r)$$

and the solution is given by $a_{jj} = e_j^T \widehat{r}$ where

$$\widehat{r} = \begin{bmatrix} H_0 & H_1 & \ldots & H_{(m-1)/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r) \text{ and}$$

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}.$$

Thus we obtain,

$$\triangle A_j = Q \begin{pmatrix} e_j^T \widehat{r} & \lambda^{m-j} \frac{(Q_1^H r)^H}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^j} Q_1^H r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \quad j = 0 : (m-1)/2.$$

Setting $X_j = 0$, we obtain $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2} \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}$. Simplifying the expressions of $\triangle A_j$s we obtain the desired result.

Next we consider spectral norm. If $|\lambda| = 1$ applying DKW Theorem 1.2.5 to $\triangle A_j$ we have $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2^2}$ and

$$X_j = -\frac{x^H r\ \lambda^{m-j} Q_1^H r (Q_1^H r)^H}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^H r|^2)},\ j = 0 : (m-1)/2.$$

This gives, $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\sqrt{m+1}} = \eta(\lambda, x, \mathrm{P})$. Simplifying the expressions for $\triangle A_j$s we obtain the desired result.

Further, for $|\lambda| \neq 1$, we consider,

$$\mu_{\triangle A_j} = \begin{cases} \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^{m-j}|^2\ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^{i}|^2\ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1 \end{cases}$$

TH-789_04612301

for $j = 0 : (m-1)/2$. Then by DKW Theorem 1.2.5 we have

$$X_j = \begin{cases} -\frac{\overline{e_j^T \widehat{r}} \ \overline{\lambda^j} \ \lambda^{m-j} \ Q_1^H r (Q_1^H r)^H}{|\lambda^{m-j}|^2 \ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ -\frac{\overline{e_j^T \widehat{r}} \ \overline{\lambda}^j \ \lambda^{m-j} \ Q_1^H r (Q_1^H r)^H}{|\lambda^i|^2 \ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1. \end{cases}$$

This gives

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{2}\sqrt{\|\widehat{r}\|^2 + \frac{\|\Pi_{(m+1)/2} R \Lambda_m\|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|^2 + \frac{\|\Pi_{(m+1)/2} \Lambda_m\|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

Simplifying the expressions of $\triangle A_j$ for all $j = 0 : (m-1)/2$ we obtain the desired result.

Now suppose that $\mathrm{P} \in \mathbb{S}$ and m is even. Notice that $A_{m/2} = A_{m/2}^H$. By Theorem 5.2.2 there exists a polynomial $\triangle \mathrm{P} \in \mathbb{S}$ such that $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. Define

$$\begin{aligned} \triangle A_j &:= Q \begin{pmatrix} a_{jj} & a_j^H \\ b_j & X_j \end{pmatrix} Q^H, \quad X_{m-j} = X_j^H, \ j = 0 : (m-2)/2, \\ \triangle A_{m/2} &:= Q \begin{pmatrix} a_{m/2,m/2} & a_{m/2}^H \\ a_{m/2} & X_{m/2} \end{pmatrix} Q^H, \quad X_{m/2} = X_{m/2}^H. \end{aligned}$$

Since $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have

$$\begin{pmatrix} \sum_{j=0}^m \lambda^j a_{jj} \\ \sum_{j=0}^{(m-2)/2} \lambda^j b_j + \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}.$$

The minimum norm solution of $\sum_{j=0}^{(m-2)/2} \lambda^j b_j + \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} = Q_1^H r$ is given by

$$b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}, \quad a_j = \frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2}, \quad a_{m/2} = \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}.$$

Note that for $|\lambda| = 1$, we have $\overline{x^H r} = \overline{\lambda^m} x^H r$. Hence the minimum norm solution of $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$ is given by $a_{jj} = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r, \ j = 0 : m$. Therefore we have

$$\begin{aligned} \triangle A_j &= Q \begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r & \frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2} (Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H, \\ \triangle A_{m/2} &= Q \begin{pmatrix} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} x^H r & \frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2} (Q_1^H r)^H \\ \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} Q_1^H r & X_{m/2} \end{pmatrix} Q^H, \\ \triangle A_{m-j} &= \triangle A_j^H, \ j = 0 : (m-2)/2, \end{aligned}$$

which gives, $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2}$. Simplifying the expressions of $\triangle A_j$ for $j = 0 : m/2$ we obtain the desired result.

Now let $|\lambda| \neq 1$. Then the value of a_{jj} from the equation $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$, $j = 0 :$

TH-789_04612301

$(m-2)/2$ is achieved after employing the condition $\overline{a}_{jj} = a_{m-j,m-j}$, $j = 0 : (m-2)/2$ and $a_{m/2,m/2} \in \mathbb{R}$ in equation (5.1) which becomes

$$\sum_{j=0}^{(m-2)/2} (\mathtt{M}(\lambda^j) + \mathtt{M}\;(\lambda^{m-j})\Sigma)\mathrm{vec}(a_{jj}) + \mathtt{M}(\lambda^{m/2})\mathrm{vec}(a_{m/2,m/2}) = \mathrm{vec}(x^H r).$$

The solution is then given by $a_{jj} = e_j^T \widehat{r}$, $j = 0 : m/2$ where

$$\widehat{r} = \begin{bmatrix} H_0 & H_1 & \ldots & H_{(m-2)/2} & H_{m/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r) \text{ and}$$

$$\begin{aligned} H_j &= \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}, \; j = 0 : (m-2)/2, \text{ and} \\ H_{m/2} &= \begin{pmatrix} \mathsf{re}\,\lambda^{m/2} \\ \mathsf{im}\,\lambda^{m/2} \end{pmatrix}. \end{aligned}$$

Therefore we obtain,

$$\begin{aligned} \triangle A_j &= Q \begin{pmatrix} e_j^T \widehat{r} & \lambda^{m-j} \frac{(Q_1^H r)^H}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^j} Q_1^H r}{\|\Lambda_m\|_2^2} & X_j \end{pmatrix} Q^H, \\ \triangle A_{m/2} &= Q \begin{pmatrix} e_{m/2}^T \widehat{r} & \lambda^{m/2} \frac{(Q_1^H r)^H}{\|\Lambda_m\|_2^2} \\ \frac{\overline{\lambda^{m/2}} Q_1^H r}{\|\Lambda_m\|_2^2} & X_{m/2} \end{pmatrix} Q^H, \\ \triangle A_{m-j} &= \triangle A_j^H, \; j = 0 : (m-2)/2. \end{aligned}$$

Setting $X_j = 0 = X_{m/2}$, $j = 0 : (m-2)/2$ we obtain

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + 2\frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expressions for $\triangle A_j$s we obtain the desired result.

Next we consider spectral norm. Note that for $|\lambda| = 1$, we have $\overline{x^H r} = \overline{\lambda^m} x^H r$. Consequently, $\lambda^j \; r^H x = \overline{\lambda^{m-j}} \; x^H r$. Consider $\mu_{\triangle A_j} = \frac{\|r\|_2}{\|\Lambda_m\|_2^2}$. Applying DKW Theorem 1.2.5 to $\triangle A_j$, we have

$$X_j = -\frac{\lambda^{m-j} \; r^H x \; Q_1^H r (Q_1^H r)^H}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |r^H x|^2)}, \; X_{m/2} = -\frac{\lambda^{m/2} \; r^H x \; Q_1^H r (Q_1^H r)^H}{\|\Lambda_m\|_2^2 \; (\|r\|_2^2 - |r^H x|^2)}$$

which gives $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{\|r\|_2}{\sqrt{m+1}}$. Simplifying the expressions for $\triangle A_j$s we obtain the desired result.

TH-789_04612301

If $|\lambda| \neq 1$ consider

$$\mu_{\triangle A_j} = \begin{cases} \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^{m-j}|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{|e_j^T \widehat{r}|^2 + \frac{|\lambda^{i}|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1 \end{cases}$$

$$\mu_{\triangle A_{m/2}} = \sqrt{|e_{m/2}^T \widehat{r}|^2 + \frac{|\lambda^m|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}$$

for $j = 0 : (m-2)/2$. Then again by DKW Theorem 1.2.5 we have

$$X_j = \begin{cases} -\frac{\overline{e_j^T \widehat{r}} \ \overline{\lambda^j} \ \lambda^{m-j} \ Q_1^H r (Q_1^H r)^H}{|\lambda^{m-j}|^2 \ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ -\frac{\overline{e_j^T \widehat{r}} \ \overline{\lambda^j} \ \lambda^{m-j} \ Q_1^H r (Q_1^H r)^H}{|\lambda^{i}|^2 \ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1, \end{cases}$$

$$X_{m/2} = -\frac{\overline{e_{m/2}^T \widehat{r}} \ Q_1^H r (Q_1^H r)^H}{(\|r\|_2^2 - |x^H r|^2)}$$

for $j = 0 : (m-2)/2$. This gives

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} R \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

Simplifying the expressions of $\triangle A_j$s we obtain the desired result.■

It is evident from the proof of Theorem 5.3.5 that, if K is a H-palindromic polynomial such that $\mathrm{P}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K}(z) = \triangle\mathrm{P}(z) + (I - xx^H)^T \mathrm{N}(z)(I - xx^H)$ for some H-palindromic polynomial N, where $\triangle$P is given in Theorem 5.3.5.

Remark 5.3.6. *If $|x^H r| = \|r\|_2$, then $\|Q_1^H r\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

Theorem 5.3.7. *Let $\mathbb{S}$ be the space of H-anti-palindromic polynomials. Assume that $\mathrm{P} \in \mathbb{S}$ and $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$. Setting $r := -\mathrm{P}(\lambda)x$, we have the following.*

1. *m be odd:*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } |\lambda| \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|\Pi_{(m+1)/2} R \Lambda_m\|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{2}\sqrt{\|\widehat{r}\|_2^2 + \frac{\|\Pi_{(m+1)/2} \Lambda_m\|^2 \ (\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

TH-789_04612301

where $\widehat{r} = \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-1)/2} \end{bmatrix}^{\dagger} \text{vec}(x^H r)$ and

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) - \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) - \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}.$$

Let $E_j := \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H - \lambda^{m-j} x r^H P_x]$. *Then for* $j = 0 : (m-1)/2$ *define*

$$\triangle A_j := \begin{cases} \dfrac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + E_j, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + E_j, & \text{if } |\lambda| \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -\triangle A_j^H$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a unique polynomial such that* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. *Next define*

$$\triangle A_j := \begin{cases} \dfrac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + E_j + \dfrac{x^H r\ \lambda^{m-j} P_x r r^H P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + E_j + \dfrac{\overline{e_j^T \widehat{r}}\ \overline{\lambda^j}\ \lambda^{m-j}\ P_x r r^H P_x}{|\lambda^{m-j}|^2\ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ e_j^T \widehat{r} x x^H + E_j + \dfrac{\overline{e_j^T \widehat{r}}\ \overline{\lambda^j}\ \lambda^{m-j}\ P_x r r^H P_x}{|\lambda^i|^2\ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1 \end{cases}$$

for $j = 0 : (m-1)/2$ *with* $\triangle A_{m-j} = -\triangle A_j^H$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a polynomial such that* $\triangle \mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle \mathrm{P}(\lambda)x = 0$ *and* $\|\triangle \mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

2. *m is even:*

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2} \leq \sqrt{2}\eta(\lambda, x, \mathrm{P}), & \text{if } |\lambda| = 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + 2\frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}, & \text{if } |\lambda| \neq 1. \end{cases}$$

$$\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \begin{cases} \eta(\lambda, x, \mathrm{P}) & \text{if } |\lambda| = 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} R \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| > 1 \\ \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + \frac{(2\|\Pi_{m/2} \Lambda_m\|_2^2 + |\lambda^{m/2}|^2)(\|r\|_2^2 - |x^H r|^2)}{\|\Lambda_m\|_2^4}}, & \text{if } |\lambda| < 1. \end{cases}$$

where $\widehat{r} = \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-2)/2} & H_{m/2} \end{bmatrix}^{\dagger} \text{vec}(x^H r)$,

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}, j = 0 : (m-2)/2, \text{ and}$$

$$H_{m/2} = \begin{pmatrix} -\mathsf{im}\,(\lambda^{m/2}) \\ \mathsf{re}\,(\lambda^{m/2}) \end{pmatrix}.$$

Let $F_j := \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^j} P_x r x^H - \lambda^{m-j} x r^H P_x]$ *and* $G_{m/2} := \frac{1}{\|\Lambda_m\|_2^2}[\overline{\lambda^{m/2}} P_x r x^H - \lambda^{m/2} x r^H P_x]$.

TH-789_04612301

Then for $j = 0 : (m-2)/2$, *define*

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x r^H x x^H + F_j, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + F_j, & \text{if } |\lambda| \neq 1 \end{cases}, \quad \triangle A_{m/2} := \begin{cases} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} x r^H x x^H + G_{m/2}, \\ \qquad \text{if } |\lambda| = 1 \\ i\, e_{m/2}^T \widehat{r} x x^H + G_{m/2}, \\ \qquad \text{if } |\lambda| \neq 1 \end{cases}$$

and $\triangle A_{m-j} = -\triangle A_j^H$. *Then* $\triangle \mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a unique polynomial such that* $\triangle\mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_F = \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$. *Next define*

$$\triangle A_j := \begin{cases} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x x^H r x^H + F_j + \frac{\lambda^{m-j}\ r^H x\ P_x r r^H P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |r^H x|^2)}, & \text{if } |\lambda| = 1 \\ e_j^T \widehat{r} x x^H + F_j + \frac{\overline{e_j^T \widehat{r}}\ \overline{\lambda^j}\ \lambda^{m-j}\ P_x r r^H P_x}{|\lambda^{m-j}|^2\ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| > 1 \\ e_j^T \widehat{r} x x^H + F_j + \frac{\overline{e_j^T \widehat{r}}\ \overline{\lambda^j}\ \lambda^{m-j}\ P_x r r^H P_x}{|\lambda^i|^2\ (\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| < 1 \end{cases}$$

$$\triangle A_{m/2} := \begin{cases} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} x x^H r x^H + G_{m/2} + \frac{\lambda^{m/2}\ r^H x\ P_x r r^H P_x}{\|\Lambda_m\|_2^2\ (\|r\|_2^2 - |r^H x|^2)}, & \text{if } |\lambda| = 1 \\ i\, e_{m/2}^T \widehat{r} x x^H + G_{m/2} + \frac{\overline{i e_{m/2}^T \widehat{r}}\ P_x r r^H P_x}{(\|r\|_2^2 - |x^H r|^2)}, & \text{if } |\lambda| \neq 1. \end{cases}$$

and $\triangle A_{m-j} = -\triangle A_j^H$ *for* $j = 0 : (m-2)/2$. *Then* $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$ *is a polynomial such that* $\triangle\mathrm{P} \in \mathbb{S}, \mathrm{P}(\lambda)x + \triangle\mathrm{P}(\lambda)x = 0$ *and* $\|\triangle\mathrm{P}\|_2 = \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

Proof: First assume that m is odd. By Theorem 5.2.2 we know that there exists $\triangle\mathrm{P} \in \mathbb{S}$, such that $\triangle\mathrm{P}(z)x + \mathrm{P}(z)x = 0$. Define

$$\triangle A_j := Q \begin{pmatrix} a_{jj} & a_j^H \\ b_j & X_j \end{pmatrix} Q^H, \quad A_{m-j} = -A_j^H, \ j = 0 : (m-1)/2.$$

Since $\triangle\mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have

$$\begin{pmatrix} \sum_{j=0}^m \lambda^j a_{jj} \\ \sum_{j=0}^{(m-1)/2} \lambda^j b_j - \sum_{j=0}^{(m-1)/2} \lambda^j a_j \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}.$$

The minimum norm solution of $\sum_{j=0}^{(m-1)/2} \lambda^j b_j - \sum_{j=0}^{(m-1)/2} \lambda^{m-j} a_j = Q_1^H r$ is given by $b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}$ and $a_j = -\frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2}$.

Note that for $|\lambda| = 1$, we have $\overline{x^H r} = -\overline{\lambda^m} x^H r$. Hence the minimum norm solution of $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$ is $a_{jj} = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r, \ j = 0 : (m-1)/2$. Thus we have

$$\begin{aligned} \triangle A_j &= Q \begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r & -\frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2} (Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H, \\ \triangle A_{m-j} &= -\triangle A_j^H \ j = 0 : (m-1)/2. \end{aligned}$$

TH-789_04612301

Setting $X_j = 0$ we have $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}} \sqrt{2\|r\|_2^2 - |r^H x|^2}$. Simplifying the expressions for $\triangle A_j$ for $j = 0 : (m-1)/2$ we obtain the desired result.

Now suppose that $|\lambda| \neq 1$. The the value of a_{jj} from $\sum_{j=0}^{m} \lambda^j a_{jj} = x^H r$, is achieved after employing the condition $\overline{a}_{jj} = -a_{m-j,m-j}$ $j = 0 : (m-1)/2$. applying vec both sides by (5.1) we have

$$\sum_{j=0}^{(m-1)/2} (\mathtt{M}(\lambda^j) - \mathtt{M}(\lambda^{m-j})\Sigma)\mathrm{vec}(a_{jj}) = \mathrm{vec}(x^H r)$$

which gives $a_{jj} = e_j^T \widehat{r}$, $j = 0 : (m-1)/2$ where $\widehat{r} = \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-1)/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r)$ and

$$H_j = \begin{pmatrix} \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) - \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) - \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}.$$

Consequently we have

$$\triangle A_j = Q \begin{pmatrix} e_j^T \widehat{r} & -\frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2}(Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r & X_j \end{pmatrix} Q^H,\ \triangle A_{m-j} = -\triangle A_j^H,\ j = 0 : (m-1)/2.$$

Setting $X_j = 0$ we obtain the minimum Frobenius norm of $\triangle A_j$ and hence we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{2} \sqrt{\|\widehat{r}\|_2^2 + \frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expression of $\triangle A_j$ we obtain the desired result.

Next consider spectral norm. From the construction of $\triangle A_j$ it is easily seen that the $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$ will be same as that given in Theorem 5.3.5 and hence desired results follow.

Now suppose that m is even. By Theorem 5.2.2 we know that there exists $\triangle \mathrm{P} \in \mathbb{S}$ such that $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$. For $j = 0 : (m-2)/2$ define

$$\triangle A_j := Q \begin{pmatrix} a_{jj} & a_j^H \\ b_j & X_j \end{pmatrix} Q^H,\ \triangle A_{m/2} = Q \begin{pmatrix} a_{m/2,m/2} & -a_{m/2}^H \\ a_{m/2} & X_{m/2} \end{pmatrix} Q^H.$$

Since $\triangle \mathrm{P}(\lambda)x + \mathrm{P}(\lambda)x = 0$, we have

$$\begin{pmatrix} \sum_{j=0}^{m} \lambda^j a_{jj} \\ \sum_{j=0}^{(m-2)/2} \lambda^j b_j - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} \end{pmatrix} = \begin{pmatrix} x^H r \\ Q_1^H r \end{pmatrix}.$$

The minimum norm solution of $\sum_{j=0}^{(m-2)/2} \lambda^j b_j - \sum_{j=0}^{(m-2)/2} \lambda^{m-j} a_j + \lambda^{m/2} a_{m/2} = Q_1^H r$ is

$$b_j = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} Q_1^H r,\ a_j = -\frac{\overline{\lambda^{m-j}}}{\|\Lambda_m\|_2^2} Q_1^H r,\ a_{m/2} = \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2} Q_1^H r.$$

Note that for $|\lambda| = 1$, we have $\overline{x^H r} = -\overline{\lambda^m} x^H r$. Hence the minimum norm solution of $\sum_{j=0}^{m} \lambda^j a_{jj} = x^H r$ is $a_{jj} = \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2} x^H r$, $j = 0 : (m-2)/2$. Note that $a_{m/2,m/2} \in i\,\mathbb{R}$. Thus we

TH-789_04612301

have

$$\begin{aligned}
\triangle A_j &= Q\begin{pmatrix} \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}x^H r & -\frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2}(Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}Q_1^H r & X_j \end{pmatrix} Q^H, \\
\triangle A_{m/2} &= Q\begin{pmatrix} \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}x^H r & -\frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2}(Q_1^H r)^H \\ \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}Q_1^H r & X_{m/2} \end{pmatrix} Q^H, \\
\triangle A_j &= -\triangle A_{m-j}^H, \; j = 0:(m-2)/2.
\end{aligned}$$

Setting $X_j = 0 = X_{m/2}$ gives $\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \frac{1}{\sqrt{m+1}}\sqrt{2\|r\|_2^2 - |r^H x|^2}$ if $|\lambda| = 1$. Simplifying the expressions for $\triangle A_j$ we obtain the desired result.

Now suppose that $|\lambda| \neq 1$. The value of a_{jj} from the equation $\sum_{j=0}^m \lambda^j a_{jj} = x^H r$, $j = 0:(m-2)/2$ is obtained after employing the condition $\overline{a}_{jj} = -a_{m-j,m-j}$ $j = 0:(m-2)/2$ and $a_{m/2,m/2} \in i\,\mathbb{R}$. Then applying vec operator both sides, by (5.1) we have $a_{jj} = e_j^T \mathrm{vec}(\widehat{r})$, $j = 0:(m-2)/2$ and $a_{m/2,m/2} = i\, e_{m/2}^T \mathrm{vec}(\widehat{r})$ where

$$\begin{aligned}
\widehat{r} &= \begin{bmatrix} H_0 & H_1 & \dots & H_{(m-2)/2} & H_{m/2} \end{bmatrix}^{\dagger} \mathrm{vec}(x^H r), \\
H_j &= \begin{pmatrix} \mathsf{re}\,(\lambda^j) + \mathsf{re}\,(\lambda^{m-j}) & -\mathsf{im}\,(\lambda^j) + \mathsf{im}\,(\lambda^{m-j}) \\ \mathsf{im}\,(\lambda^j) + \mathsf{im}(\lambda^{m-j}) & \mathsf{re}\,(\lambda^j) - \mathsf{re}\,(\lambda^{m-j}) \end{pmatrix}, \; j = 0:(m-2)/2, \text{ and} \\
H_{m/2} &= \begin{pmatrix} -\mathsf{im}\,(\lambda^{m/2}) \\ \mathsf{re}\,(\lambda^{m/2}) \end{pmatrix}.
\end{aligned}$$

Consequently we have

$$\begin{aligned}
\triangle A_j &= Q\begin{pmatrix} e_j^T \widehat{r} & -\frac{\lambda^{m-j}}{\|\Lambda_m\|_2^2}(Q_1^H r)^H \\ \frac{\overline{\lambda^j}}{\|\Lambda_m\|_2^2}Q_1^H r & X_j \end{pmatrix} Q^H, \\
\triangle A_{m/2} &= Q\begin{pmatrix} i\, e_{m/2}^T \widehat{r} & -\frac{\lambda^{m/2}}{\|\Lambda_m\|_2^2}(Q_1^H r)^H \\ \frac{\overline{\lambda^{m/2}}}{\|\Lambda_m\|_2^2}Q_1^H r & X_{m/2} \end{pmatrix} Q^H, \\
\triangle A_j &= -\triangle A_{m-j}^H, \; j = 0:(m-2)/2.
\end{aligned}$$

Setting $X_j = 0 = X_{m/2}$ we obtain

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P}) = \sqrt{(2\|\widehat{r}\|_2^2 - |e_{m/2}^T \widehat{r}|^2) + 2\frac{\|r\|_2^2 - |x^H r|^2}{\|\Lambda_m\|_2^2}}.$$

Simplifying the expressions for $\triangle A_j$'s we obtain the desired result.

Next consider spectral norm. From the construction of $\triangle A_j$ it is easily seen that $\eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$ will be same as that given in Theorem 5.3.5 and hence the proof is similar.■

It follows from the proof of Theorem 5.3.5 that, if K is a H-anti-palindromic polynomial

TH-789_04612301

such that $\mathrm{P}(\lambda)x + \mathrm{K}(\lambda)x = 0$ then $\mathrm{K}(z) = \triangle\mathrm{P}(z) + (I - xx^H)^T\mathrm{N}(z)(I - xx^H)$ for some H-anti-palindromic polynomial N, where $\triangle$P is given in Theorem 5.3.7.

Remark 5.3.8. *If $|x^H r| = \|r\|_2$, then $\|Q_1^H r\|_2 = 0$. In such a case, considering $X_j = 0$, $j = 0 : m$, we obtain the desired results.*

5.4 Structured backward error and palindromic linearizations

As in the case of structured polynomials discussed in Chapter 4, the task of choosing a "good" linearization is crucial for palindromic polynomials as well. It is shown in [68] by Mackey et al. that, a $*$-palindromic or $*$-anti-palindromic polynomial P can have both $*$-palindromic and $*$-anti-palindromic linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$, for $* = T/H$ that preserve the eigen-symmetry of P. To be specific, for $* = H$ there always exist a structured linearization and for $* = T$, it does only when either $1 \notin \sigma(\mathrm{P})$ or $-1 \notin \sigma(\mathrm{P})$.

Since plenty of structured linearizations can be constructed for a given $\mathrm{P} \in \mathbb{S}$, it is necessary to identify "good" linearizations. For the rest of the section we set $\Lambda_{m-1} := [\lambda^{m-1}, \lambda^{m-2}, \ldots, \lambda, 1]^T$.

Recall that $x \in \mathbb{C}^n$ is a right eigenvector of P corresponding to an eigenvalue λ if and only if $\Lambda_{m-1} \otimes x$ is a right eigenvector of $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ corresponding to the eigenvalue λ. Also recall that

$$\|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2 = \|v\|_2 \|\mathrm{P}(\lambda)x\|_2 \text{ and } |(\Lambda_{m-1} \otimes x)^T \mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)| = |\Lambda_{m-1}^T v|\, |x^T \mathrm{P}(\lambda)x|$$

whenever $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is the linearization corresponding to an ansatz vector $v \in \mathbb{C}^m$. Without loss of generality we assume that the ansatz vector $v \in \mathbb{C}^m$ is of unit norm, that is, $\|v\|_2 = 1$.

Assume that (λ, x) is an approximate eigenpair of P. Then by Theorem 4.4.1 we have

$$\frac{1}{\sqrt{2}} \leq \frac{\eta(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$$

for both $\|\cdot\| \equiv \|\cdot\|_F$ and $\|\cdot\| \equiv \|\cdot\|_2$, where $(\lambda, \Lambda_{m-1} \otimes x)$ is an approximate eigenpair of the linearization L associated with the ansatz vector $v \in \mathbb{C}^m$.

Table 5.2 gives the structured of ansatz vectors for structured polynomials, see [64, 68].

Let $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ be the structured linearization of $\mathrm{P} \in \mathbb{S}$ corresponding to an ansatz vector v. Then by Lemma 4.4.2 we have

$$\frac{\eta^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \geq \frac{1}{\sqrt{2}}. \tag{5.2}$$

Also recall that

$$\frac{1}{\sqrt{2}} \leq \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2 \, \|(1, \lambda)\|_2} \leq 1. \tag{5.3}$$

Set $\delta_v := \dfrac{|\Lambda_{m-1}^* v|}{\|\Lambda_{m-1}\|_2}, * \in \{T, H\}$. Then $0 < \delta_v^{-1} \leq 1$.

Theorem 5.4.1. *Let $\mathbb{S}$ be the space of T-palindromic matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$. Let $\mathrm{L}_p \in \mathbb{L}_1(\mathrm{P})$ be a T-palindromic linearization and $\mathrm{L}_a \in \mathbb{L}_1(\mathrm{P})$ be a T-anti-palindromic*

TH-789_04612301

$\mathbb{S}$	Structured Linearization	ansatz vector
T-palindromic	T-palindromic	$Rv = v$
	T-anti-palindromic	$Rv = -v$
T-anti-palindromic	T-palindromic	$Rv = -v$
	T-anti-palindromic	$Rv = v$
H-palindromic	H-palindromic	$Rv = \overline{v}$
	H-anti-palindromic	$Rv = -\overline{v}$
H-anti-palindromic	H-palindromic	$Rv = -\overline{v}$
	H-anti-palindromic	$Rv = \overline{v}$

Table 5.2: Table for the admissible ansatz vectors for palindromic polynomials.

linearization of P *with respect to the ansatz vectors $Rv = v$ and $Rv = -v$ respectively. If (λ, x) is an approximate right eigenpair of* P *then we have the following:*

1. $\|\cdot\| \equiv \|\cdot\|_F : 1 \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ *if* $\lambda = -1$

$$\sqrt{1 - 2\,\mathsf{re}(\lambda)|1+\lambda|^{-2}\delta_v^{-2}} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2} \text{ if } \mathsf{re}(\lambda) \geq 0$$

$$1 \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{1 - 2\,\mathsf{re}(\lambda)|1+\lambda|^{-2}\delta_v^{-2}} \text{ if } \mathsf{re}(\lambda) < 0, \lambda \neq -1$$

$$1 \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_a; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}, \text{ if } \lambda = 1$$

$$1 \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_a; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\sqrt{1 + 2\,\mathsf{re}(\lambda)|1-\lambda|^{-2}\delta_v^{-2}} \text{ if } \mathsf{re}(\lambda) \geq 0, \lambda \neq 1$$

$$\sqrt{1 + 2\,\mathsf{re}(\lambda)|1-\lambda|^{-2}\delta_v^{-2}} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_a; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2} \text{ if } \mathsf{re}(\lambda) < 0$$

2. $\|\cdot\| \equiv \|\cdot\|_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$ *if* $\lambda = \pm 1$

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\,\sqrt{1 + |1+\lambda|^{-2}\delta_v^{-2}}\;\eta(\lambda, x, \mathrm{P}) \text{ if } \lambda \neq \pm 1$$

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_a; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1 \text{ if } \lambda = \pm 1$$

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_a; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}\,\sqrt{1 + |1-\lambda|^{-2}\delta_v^{-2}}\;\eta(\lambda, x, \mathrm{P}) \text{ if } \lambda \neq \pm 1.$$

Proof: First consider L_p. If $\lambda = -1$ then by Theorem 5.3.1 we have

$$\begin{aligned}
\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v) &= \sqrt{2}\frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2}{\|(\Lambda_{m-1} \otimes x)\|_2\,\|(1,\ \lambda)\|_2} = \sqrt{2}\frac{\|\mathrm{P}(\lambda)x\|_2}{\|(\Lambda_{m-1} \otimes x)\|_2\,\|(1,\ \lambda)\|_2} \\
&= \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2\|x\|_2\|(1,\ \lambda)\|_2}\,\eta(\lambda, x, \mathrm{P}) \text{ and} \\
\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}_p; v) &= \frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes x)\|_2}{\|(\Lambda_{m-1} \otimes x)\|_2\,\|(1,\ \lambda)\|_2} = \frac{\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2\|x\|_2\|(1,\ \lambda)\|_2}\,\eta(\lambda, x, \mathrm{P}).
\end{aligned}$$

TH-789_04612301

Therefore by (5.3) the desired result follows for $\lambda = -1$. If $\lambda \neq -1$ then for the $|||\cdot||| \equiv |||\cdot|||_F$ we have

$$\eta_F^{\mathbb{S}}(\lambda, x, \mathrm{L}_p; v) = \sqrt{2}\,\frac{\sqrt{\frac{\|\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2} - 2\frac{\mathsf{re}(\lambda)}{|1+\lambda|^2}\frac{|(\Lambda_{m-1}\otimes x)^T\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)|^2}{\|(\Lambda_{m-1}\otimes x)\|_2^4}}}{\|(1,\,\lambda)\|_2}$$

$$= \sqrt{2}\,\frac{\sqrt{\|\mathrm{P}(\lambda)x\|_2^2 - 2\frac{\mathsf{re}(\lambda)}{|1+\lambda|^2}\frac{|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}|x^T\mathrm{P}(\lambda)x|^2}}{\|(1,\,\lambda)\|_2\,\|\Lambda_{m-1}\|_2}.$$

The desired result follows by using $|x^T\mathrm{P}(\lambda)x| \leq \|\mathrm{P}(\lambda)x\|_2$ and (5.3) for both $\mathsf{re}(\lambda) \geq 0$ and $\mathsf{re}(\lambda) < 0$. Now consider $|||\cdot||| \equiv |||\cdot|||_2$. Then the desired result is immediate by Theorem 5.3.1 and (5.3) for $\lambda = \pm 1$. Further for $\lambda \neq \pm 1$, by Theorem 5.3.1 we have the following.

If $|\lambda| > 1$ then

$$\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_p; v) \leq \sqrt{2}\sqrt{\frac{|(\Lambda_{m-1}\otimes x)^T\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)|^2}{|1+\lambda|^2\,\|(\Lambda_{m-1}\otimes x)\|_2^4} + \frac{|\lambda|^2\,\|\mathrm{L}(\lambda)(\Lambda_{m-1}\otimes x)\|_2^2}{\|(\Lambda_{m-1}\otimes x)\|_2^2\,\|(1,\,\lambda)\|_2^4}}$$

$$= \sqrt{2}\sqrt{\frac{|\Lambda_{m-1}^T v|^2\,|x^T\mathrm{P}(\lambda)x|^2}{|1+\lambda|^2\,\|\Lambda_{m-1}\|_2^4} + \frac{|\lambda|^2\,\|r\|_2^2}{\|\Lambda_{m-1}\|_2^2\,\|(1,\,\lambda)\|_2^4}}$$

$$\leq \frac{\sqrt{2}\,\|r\|_2}{\|\Lambda_{m-1}\|_2\,\|(1,\,\lambda)\|_2}\sqrt{\frac{|\Lambda_{m-1}^T v|^2}{|1+\lambda|^2\|\Lambda_{m-1}\|_2^2} + \frac{|\lambda|^2}{\|(1,\,\lambda)\|_2^2}}$$

Notice that $|\lambda|^2 \leq \|(1,\,\lambda)\|_2^2$. Hence by (5.3) and (5.2) the desired result follows. Similarly for $|\lambda| \leq 1$, by Theorem 5.3.1 we have

$$\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_p; v) \leq \sqrt{2}\sqrt{\frac{|\Lambda_{m-1}^T v|^2}{|1+\lambda|^2\|\Lambda_{m-1}\|_2^2} + \frac{1}{\|(1,\,\lambda)\|_2^2}}\;\eta(\lambda, x, \mathrm{P}).$$

Hence the result follows by (5.2).

Next consider L_a. If $\lambda = 1$ then results follow by Theorem 5.3.3 and (5.3). If $\lambda \neq 1$ then for the $|||\cdot||| \equiv |||\cdot|||_F$ we have

$$\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}_a; v) = \sqrt{2}\,\frac{\sqrt{\|\mathrm{P}(\lambda)x\|_2^2 + 2\frac{\mathsf{re}(\lambda)}{|1-\lambda|^2}\frac{|\Lambda_{m-1}^T v|^2}{\|\Lambda_{m-1}\|_2^2}|x^T\mathrm{P}(\lambda)x|^2}}{\|(1,\,\lambda)\|_2\,\|\Lambda_{m-1}\|_2}.$$

Using $|x^T\mathrm{P}(\lambda)x| \leq \|\mathrm{P}(\lambda)x\|_2$ and (5.3) the desired result follows for both $\mathsf{re}(\lambda) \geq 0$ and $\mathsf{re}(\lambda) < 0$. The proof is similar for $|||\cdot||| \equiv |||\cdot|||_2$.■

We mention that bounds in Theorem 5.4.1 hold when P is T-anti-palindromic.

The bounds obtained in the Theorem 5.4.1 advise that, choose T-palindromic linearization for $\mathsf{re}(\lambda) \geq 0$ and choose T-anti-palindromic linearization for $\mathsf{re}(\lambda) < 0$.

Theorem 5.4.2. *Let $\mathbb{S}$ be the space of H-palindromic matrix polynomials. Assume that $\mathrm{P} \in \mathbb{S}$. Suppose $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ is the H-palindromic linearization or H-anti-palindromic linearization of P corresponding to an ansatz vector v. If (λ, x) is an approximate right eigenpair of P then we have the following:*

1. $|||\cdot||| \equiv |||\cdot|||_F$: $\dfrac{\sqrt{2-\delta_v^{-2}}}{\sqrt{2}} \leq \dfrac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \sqrt{2}$ *if* $|\lambda| = 1$

TH-789_04612301

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \|(1,\, \lambda)\|_2^{-2}} \text{ if } |\lambda| \neq 1 \text{ where}$$

$$\tau = \left\| \begin{bmatrix} 1+\mathsf{re}(\lambda) & \mathsf{im}(\lambda) \\ \mathsf{im}(\lambda) & 1-\mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2 \text{ for } H\text{-palindromic linearization and}$$

$$\tau = \left\| \begin{bmatrix} 1-\mathsf{re}(\lambda) & -\mathsf{im}(\lambda) \\ -\mathsf{im}(\lambda) & 1+\mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2 \text{ for } H\text{-anti-palindromic linearization.}$$

2. $|||\cdot||| \equiv |||\cdot|||_2 : \dfrac{1}{\sqrt{2}} \leq \dfrac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq 1$ *if* $|\lambda| = 1$

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \frac{|\lambda|^2}{\|(1,\, \lambda)\|_2^4}} \text{ if } |\lambda| > 1$$

$$\frac{1}{\sqrt{2}} \leq \frac{\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v)}{\eta(\lambda, x, \mathrm{P})} \leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \frac{1}{\|(1,\, \lambda)\|_2^4}} \text{ if } |\lambda| < 1$$

Proof: Assume that L is an H-palindromic. If $|\lambda| = 1$, by Theorem 5.3.5 and Theorem 5.3.7, for the $|||\cdot||| \equiv |||\cdot|||_F$ we have,

$$\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) = \frac{1}{\|(1,\, \lambda)\|_2\, \|\Lambda_{m-1}\|_2}\sqrt{2\|\mathrm{P}(\lambda)x\|_2^2 - \frac{|\Lambda_{m-1}^H v|^2\, |x^H \mathrm{P}(\lambda)x|^2}{\|\Lambda_{m-1}\|_2^2}}.$$

Using $|x^H\mathrm{P}(\lambda)x| \leq \|\mathrm{P}(\lambda)x\|_2$ and (5.3) the result follows. The proof is similar for the case when $|||\cdot||| \equiv |||\cdot|||_2$. Next, let $|\lambda| \neq 1$. Then by Theorem 5.3.5 and Theorem 5.3.7 we have

$$\begin{aligned} \eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) &= \sqrt{2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^4} + \frac{\frac{\|\mathrm{P}(\lambda)x\|_2^2}{\|\Lambda_{m-1}\|_2^2} - \frac{|\Lambda_{m-1}^H v|^2}{\|\Lambda_{m-1}\|_2^4}|x^H\mathrm{P}(\lambda)x|^2}{\|(1,\, \lambda)\|_2^2}} \\ &\leq \frac{\sqrt{2}}{\|\Lambda_{m-1}\|_2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^2} + \frac{\|r\|_2^2}{\|(1,\, \lambda)\|_2^2}} \end{aligned}$$

for $|||\cdot||| \equiv |||\cdot|||_F$, where $\widehat{r} = \begin{bmatrix} 1+\mathsf{re}(\lambda) & \mathsf{im}(\lambda) \\ \mathsf{im}(\lambda) & 1-\mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \begin{bmatrix} \mathsf{re}(\Lambda_{m-1}^H v\, x^H\mathrm{P}(\lambda)x) \\ \mathsf{im}(\Lambda_{m-1}^H v\, x^H\mathrm{P}(\lambda)x) \end{bmatrix}$.

Notice that $\|\widehat{r}\|_2 \leq \tau|\Lambda_{m-1}^H v|\|\mathrm{P}(\lambda)x\|_2$ where $\tau = \left\| \begin{bmatrix} 1+\mathsf{re}(\lambda) & \mathsf{im}(\lambda) \\ \mathsf{im}(\lambda) & 1-\mathsf{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2$. Thus we obtain

$$\eta_F^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) \leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \frac{1}{\|(1,\, \lambda)\|_2^2}}\,\eta(\lambda, x, \mathrm{P}).$$

Hence the result follows. Next consider $|||\cdot||| \equiv |||\cdot|||_2$ and $|\lambda| \neq 1$. Then we have the following.

If $|\lambda| > 1$ then by Theorem 5.3.5 and Theorem 5.3.7 we have

$$\begin{aligned} \eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1}\otimes x, \mathrm{L}; v) &= \sqrt{2}\sqrt{\frac{\|\widehat{r}\|_2^2}{\|\Lambda_{m-1}\|_2^4} + \frac{|\lambda|^2\, \|r\|_2^2}{\|\Lambda_{m-1}\|_2^2\, \|(1, \lambda)\|_2^4}} \\ &\leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \frac{|\lambda|^2}{\|(1,\, \lambda)\|_2^4}}\,\eta(\lambda, x, \mathrm{P}). \end{aligned}$$

TH-789_04612301

Similarly, if $|\lambda| < 1$, we have

$$\eta_2^{\mathbb{S}}(\lambda, \Lambda_{m-1} \otimes x, \mathrm{L}; v) \leq \frac{\sqrt{2}\|\Lambda_m\|_2}{\|\Lambda_{m-1}\|_2}\sqrt{\tau^2\delta_v^{-2} + \frac{1}{\|(1,\, \lambda)\|_2^4}}\eta(\lambda, x, \mathrm{P}).$$

Hence the result follows by (5.2). The proof is similar when L is an H-anti-palindromic linearization. Hence the proof.■

The morale of the Theorem 5.4.2 is as follows. For $|\lambda| = 1$, it does not really make any difference in the magnified backward error due to either H-palindromic or H-anti-palindromic linearization and hence any of H-palindromic or H-anti-palindromic linearization can be chosen. For $|\lambda| \neq 1$, two cases arise. Let us define $\tau := \tau_p$ for H-palindromic linearization and $\tau := \tau_a$ for H-anti-palindromic linearization. Then for a given (λ, x) if $\tau_a \leq \tau_a$ then choose H-palindromic linearization else choose H-anti-palindromic linearization.

5.5 Pseudospectra of palindromic polynomials

Now we consider structured backward error of an approximate eigenvalue of $\mathrm{P} \in \mathbb{S}$. We have seen in chapter 4, that this could be applied gainfully to obtain partial equality between structured and unstructured pseudospectra. Table 5.3 that illustrates the following: for $\mathrm{P} \in \mathbb{S}$ there always exist certain approximate eigenvalues for which $\eta^{\mathbb{S}}(\lambda, x, \mathrm{P}) \leq \alpha\eta(\lambda, x, \mathrm{P})$ for some $\alpha > 0$. Recall that $\eta(\lambda, \mathrm{P}) = \min_{\|x\|=1} \eta(\lambda, x, \mathrm{P})$, $\eta_F^{\mathbb{S}}(\lambda, \mathrm{P}) = \min_{\|x\|=1} \eta_F^{\mathbb{S}}(\lambda, x, \mathrm{P})$ and $\eta_2^{\mathbb{S}}(\lambda, \mathrm{P}) = \min_{\|x\|=1} \eta_2^{\mathbb{S}}(\lambda, x, \mathrm{P})$.

$\mathbb{S}$	m	$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) \leq \alpha\eta(\lambda, \mathrm{P})$ $\|\cdot\| \equiv \|\cdot\|_F$	$\eta^{\mathbb{S}}(\lambda, \mathrm{P}) \leq \alpha\eta(\lambda, \mathrm{P})$ $\|\cdot\| \equiv \|\cdot\|_2$
T-palindromic	odd	$\alpha = \sqrt{2}, \lambda = -1$	$\alpha = 1, \lambda = \pm 1$
	even	–	$\alpha = 1, \lambda = \pm 1$
T-anti-palindromic	odd	$\alpha = \sqrt{2}, \lambda = 1$	$\alpha = 1, \lambda = 1$
	even	$\alpha = \sqrt{2}, \lambda = 1$	$\alpha = 1, \lambda = 1$
H-palindromic/ H-anti-palindromic	odd/even	$\alpha = \sqrt{2}, \lvert\lambda\rvert = 1$	$\alpha = 1, \lvert\lambda\rvert = 1$

Table 5.3: Partial equality of structured backward errors

Theorem 5.5.1. *Let $\mathbb{S}$ be the space of H-palindromic or H-antipalindromic matrix polynomials of degree m and $\mathrm{P} \in \mathbb{S}$. Let $\lambda \in \mathbb{T}$. Then for the spectral and the Frobenius norms on $\mathbb{C}^{n\times n}$, we have $\eta^{\mathbb{S}}(\lambda, \mathrm{P}) = \eta(\lambda, \mathrm{P})$ and hence $\sigma_\epsilon^{\mathbb{S}}(\mathrm{P}) \cap \mathbb{T} = \sigma_\epsilon(\mathrm{P}) \cap \mathbb{T}$.*

Proof: First assume that $\mathrm{P} \in \mathbb{P}_m(\mathbb{C}^{n\times n})$ be an H-palindromic polynomial. Let $\lambda \in \mathbb{T}$. We have $\mathrm{P}(\lambda)^H = \overline{\lambda^m}\mathrm{P}(\lambda)$. This shows that $\mathrm{P}(\lambda)$ is a normal matrix. Let (μ, u) be an eigenpair of $\overline{\lambda^m}\mathrm{P}(\lambda)$ such that $|\mu| = \sigma_{\min}(\overline{\lambda}^m\mathrm{P}(\lambda)) = \sigma_{\min}(\mathrm{L}(\lambda))$. Define $\triangle A_j := -\frac{1}{2}\lambda^j\mu\, uu^H$ and consider the polynomial $\triangle\mathrm{P}(z) = \sum_{j=0}^m z^j \triangle A_j$. Noting the fact that $\overline{\lambda^m}\mathrm{P}(\lambda)u = \mu u$ and

TH-789_04612301

$\overline{\mu}u = (\overline{\lambda^m}\mathrm{P}(\lambda))^H u = \mu u$, we have $\mathrm{P}(\lambda)u + \triangle\mathrm{P}(\lambda)u = \lambda^m \mu\, u - \lambda^m \mu u = 0$. Further, we have

$$\|\triangle\mathrm{P}\| = \frac{\sigma_{\min}(\mathrm{P}(\lambda))}{\|\Lambda_m\|_2} = \eta(\lambda, \mathrm{P}).$$

Hence the results follow. ■

Indian Institute of Technology Guwahati

TH-789_04612301

structured polynomial	eigenvalue of interest	choice of linearization	$\beta^{\mathbb{S}} := \eta^S(\lambda, x, \mathrm{L})/\eta(\lambda, \Lambda_{m-1}\otimes x, \mathrm{P})$ $\lVert\cdot\rVert \equiv \lVert\cdot\rVert_F$	$\beta^{\mathbb{S}} := \eta^S(\lambda, x, \mathrm{L})/\eta(\lambda, \Lambda_{m-1}\otimes x, \mathrm{P})$ $\lVert\cdot\rVert \equiv \lVert\cdot\rVert_2$
T-palindromic/ T-anti-palindromic	$\mathrm{re}(\lambda) \geq 0$	T-palindromic	$\sqrt{1 - 2\mathrm{re}(\lambda)\|1+\lambda\|^{-2}\delta_v^{-2}} \leq \alpha^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1 + \|1+\lambda\|^{-2}\delta_v^{-2}}$
	$\mathrm{re}(\lambda) \leq 0$	T-anti-palindromic	$\sqrt{1 + 2\mathrm{re}(\lambda)\|1-\lambda\|^{-2}\delta_v^{-2}} \leq \alpha^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}\sqrt{1 + \|1-\lambda\|^{-2}\delta_v^{-2}}$
H-palindromic/ H-anti-palindromic	$\|\lambda\| = 1$	H-palindromic/ H-anti-palindromic	$\frac{\sqrt{2-\delta_v^{-2}}}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq \sqrt{2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 1$
	$\|\lambda\| \neq 1$	H-palindromic	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{1 + \tau_p^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{\lVert(1,\lambda)\rVert_2^{-2} + \tau_p^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$ if $\|\lambda\| < 1$
				$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{\lVert(1,\lambda^{-1})\rVert_2^{-2} + \tau_p^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$ if $\|\lambda\| > 1$
		H-anti-palindromic	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{1 + \tau_a^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$	$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{\lVert(1,\lambda)\rVert_2^{-2} + \tau_a^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$ if $\|\lambda\| < 1$
				$\frac{1}{\sqrt{2}} \leq \beta^{\mathbb{S}} \leq 2\sqrt{\lVert(1,\lambda^{-1})\rVert_2^{-2} + \tau_a^2\delta_v^{-2}\lVert(1,\lambda)\rVert_2^2}$ if $\|\lambda\| > 1$

Table 5.4: The choice of structured linearizations for palindromic polynomials, where $\delta_v = \|\Lambda_{m-1}\|_2/|\Lambda_{m-1}^* v|$, $* \in \{T, H\}$ and v is the ansatz vector.

$$\text{Here } \tau_p := \left\| \begin{bmatrix} 1+\mathrm{re}(\lambda) & \mathrm{im}(\lambda) \\ \mathrm{im}(\lambda) & 1-\mathrm{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2 \text{ and } \tau_a := \left\| \begin{bmatrix} 1-\mathrm{re}(\lambda) & -\mathrm{im}(\lambda) \\ -\mathrm{im}(\lambda) & 1+\mathrm{re}(\lambda) \end{bmatrix}^{\dagger} \right\|_2 .$$

TH-789_04612301

Chapter 6

Structured eigenvalue condition numbers and linearizations for matrix polynomials

This chapter is concerned with the sensitivity analysis of eigenvalue problems for structured matrix polynomials, including complex symmetric, Hermitian, even, odd, palindromic, and anti-palindromic matrix polynomials. As mentioned before, numerical methods for solving such eigenvalue problem proceed by linearizing the matrix polynomial into a matrix pencil of larger size. A question of practical importance is whether this process of linearization increases the sensitivity of the eigenvalue with respect to structured perturbations. For all structures under consideration, we show that this is not the case: there is always a linearization for which the structured condition number of an eigenvalue does not differ significantly. This implies, for example, that a structure-preserving algorithm applied to the linearization fully benefits from a potentially low structured eigenvalue condition number of the original matrix polynomial.

6.1 Introduction

Consider an $n \times n$ matrix polynomial

$$\mathrm{P}(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2 + \cdots + \lambda^m A_m, \tag{6.1}$$

with $A_0, \ldots, A_m \in \mathbb{C}^{n\times n}$. An eigenvalue $\lambda \in \mathbb{C}$ of P, defined by the relation $\det(\mathrm{P}(\lambda)) = 0$, is called simple if λ is a simple root of the polynomial $\det(\mathrm{P}(\lambda))$.

This chapter is concerned with the sensitivity of a simple eigenvalue λ under perturbations of the coefficients A_j. The condition number of λ is a first-order measure for the worst-case effect of perturbations on λ. Tisseur [93] has provided an explicit expressions for this condition number. Subsequently, this expression was extended to polynomials in homogeneous form by Dedieu and Tisseur [26], see also [1, 14, 21], and to semi-simple eigenvalues in [54]. In the more general context of nonlinear eigenvalue problems, the sensitivity of eigenvalues and eigenvectors has been investigated in, e.g., [4, 58, 60, 61]. We consider the same classes of structured polynomials as discussed in Table 1.2.

TH-789_04612301

In certain situations, it is reasonable to expect that perturbations of the polynomial respect the underlying structure. For example, if a strongly backward stable eigenvalue solver is applied to a palindromic matrix polynomial then the computed eigenvalues would be the exact eigenvalues of a slightly perturbed *palindromic* eigenvalue problems. Also, structure-preserving perturbations are physically more meaningful in the sense that the spectral symmetries induced by the structure are not destroyed. Restricting the admissible perturbations might have a positive effect on the sensitivity of an eigenvalue. This question has been studied for linear eigenvalue problems in quite some detail recently [19, 38, 49, 50, 51, 54, 73, 78, 81]. It often turns out that the desirable positive effect is not very remarkable: in many cases the worst-case eigenvalue sensitivity changes little or not at all when imposing structure. Notable exceptions can be found among symplectic, skew-symmetric, and palindromic eigenvalue problems [51, 54]. In the first part of this chapter, we will extend these results to structured matrix polynomials.

Due to the lack of a robust genuine polynomial eigenvalue solver, the eigenvalues of P are usually computed by first reformulating (6.1) as an $mn \times mn$ linear generalized eigenvalue problem and then applying a standard method such as the QZ algorithm [34] to the linear problem. This process of linearization introduces unwanted effects. Besides the obvious increase of dimension, it may also happen that the eigenvalue sensitivities significantly deteriorate. Fortunately, one can use the freedom in the choice of linearization to minimize this deterioration for the eigenvalue region of interest, as proposed for quadratic eigenvalue problems in [29, 42, 93]. For the general polynomial eigenvalue problem (6.1), Higham et al. [39, 41] have identified linearizations with minimal eigenvalue condition number/backward error among the set of linearizations described in [67]. For structured polynomial eigenvalue problem, rather than using *any* linearization it is of course advisable to use one which has a similar structure. For example, it was shown in [68] that a palindromic matrix polynomial can usually be linearized into a palindromic or anti-palindromic matrix pencil, offering the possibility to apply structure-preserving algorithms to the linearization. It is natural to ask whether there is also a structured linearization that has no adverse effect on the structured condition number. For a small subset of structures from Table 1.2, this question has already been discussed in [42]. In the second part of this chapter, we extend the discussion to all structures from Table 1.2.

The rest of this chapter is organized as follows. In Section 6.2, we first review the derivation of the unstructured eigenvalue condition number for a matrix polynomial and then provide explicit expressions for structured eigenvalue conditions numbers. Most but not all of these expressions are generalizations of known results for linear eigenvalue problems. In Section 6.4, we apply these results to find good choices from the set of structured linearizations described in [68].

6.2 Structured condition numbers for matrix polynomials

Before discussing the effect of structure on the sensitivity of an eigenvalue, we briefly review existing results on eigenvalue condition numbers for matrix polynomials. Assume that λ is a *simple finite* eigenvalue of the matrix polynomial P defined in (6.1) with normalized right

TH-789_04612301

and left eigenvectors x and y:

$$\mathrm{P}(\lambda)x = 0, \quad y^H \mathrm{P}(\lambda) = 0, \quad \|x\|_2 = \|y\|_2 = 1. \tag{6.2}$$

The perturbation

$$(\mathrm{P} + \triangle\mathrm{P})(\lambda) = (A_0 + E_0) + \lambda(A_1 + E_1) + \cdots + \lambda^m (A_m + E_m)$$

moves λ to an eigenvalue $\widehat{\lambda}$ of $\mathrm{P} + \triangle\mathrm{P}$. A useful tool to study the effect of $\triangle\mathrm{P}$ is the first order *perturbation expansion*

$$\widehat{\lambda} = \lambda - \frac{1}{y^H \mathrm{P}'(\lambda)x} y^H \triangle\mathrm{P}(\lambda)x + O(\|\triangle\mathrm{P}\|^2), \tag{6.3}$$

which can be derived, e.g., by applying the implicit function theorem to (6.2), see [26, 95]. Note that $y^H \mathrm{P}'(\lambda)x \neq 0$ because λ is simple [2, 4].

To measure the sensitivity of λ we first need to specify a way to measure $\triangle\mathrm{P}$. Given a matrix norm $\|\cdot\|_M$ on $\mathbb{C}^{n\times n}$, a monotone vector norm $\|\cdot\|_v$ on $\mathbb{C}^{m+1}$ and non-negative weights $\omega_0, \ldots, \omega_m$, we define

$$\|\triangle\mathrm{P}\| := \left\| \left[\frac{1}{\omega_0}\|E_0\|_M, \ \frac{1}{\omega_1}\|E_1\|_M, \ \ldots, \ \frac{1}{\omega_m}\|E_m\|_M \right] \right\|_v . \tag{6.4}$$

A relatively small weight ω_j means that $\|E_j\|_M$ will be small compared to $\|\triangle\mathrm{P}\|$. In the extreme case $\omega_j = 0$, we define $\|E_j\|_M/\omega_j = 0$ for $\|E_j\|_M = 0$ and $\|E_j\|_M/\omega_j = \infty$ otherwise. If all ω_j are positive then (6.4) defines a norm on $\mathbb{C}^{n\times n} \times \cdots \times \mathbb{C}^{n\times n}$. See [1, 2] for more on norms of matrix polynomials.

We are now ready to introduce a condition number for the eigenvalue λ of P with respect to the choice of $\|\triangle\mathrm{P}\|$ in (6.4):

$$\kappa_{\mathrm{P}}(\lambda) := \lim_{\epsilon \to 0} \sup \left\{ \frac{|\hat{\lambda} - \lambda|}{\epsilon} : \ \|\triangle\mathrm{P}\| \leq \epsilon \right\}, \tag{6.5}$$

where $\hat{\lambda}$ is the eigenvalue of $\mathrm{P} + \triangle\mathrm{P}$ closest to λ. An explicit expression for $\kappa_{\mathrm{P}}(\lambda)$ can be found in [95, Thm. 5] for the case $\|\cdot\|_v \equiv \|\cdot\|_\infty$ and $\|\cdot\|_M \equiv \|\cdot\|_2$. In contrast, the approach used in [26] requires an accessible geometry on the perturbation space and thus facilitates the norms $\|\cdot\|_v \equiv \|\cdot\|_2$ and $\|\cdot\|_M \equiv \|\cdot\|_F$. On the other hand, the approach used in [1] analyzes condition number by considering any norm $\|\cdot\|_M$ and the Hölder's p-norm $\|\cdot\|_V \equiv \|\cdot\|_p$. The following lemma includes both settings. Note that the dual to the vector norm $\|\cdot\|_v$ is defined as

$$\|w\|_d := \sup_{\|z\|_v \leq 1} |w^T z|,$$

see, e.g., [36].

Lemma 6.2.1. *Consider the condition number $\kappa_{\mathrm{P}}(\lambda)$ defined in (6.5) with respect to the semi-norm (6.4). For any unitarily invariant norm $\|\cdot\|_M$ we have*

$$\kappa_{\mathrm{P}}(\lambda) = \frac{\|[\omega_0, \ \omega_1|\lambda|, \ \ldots, \omega_m|\lambda|^m]\|_d}{|y^H \mathrm{P}'(\lambda)x|} \tag{6.6}$$

TH-789_04612301

where $\|\cdot\|_d$ *denotes the vector norm dual to* $\|\cdot\|_v$.

Proof: Inserting the perturbation expansion (6.3) into (6.5) yields

$$\kappa_{\mathrm{P}}(\lambda) = \frac{1}{|y^H \mathrm{P}'(\lambda)x|} \sup\left\{|y^H \triangle\mathrm{P}(\lambda)x| : \ \|\triangle\mathrm{P}\| \leq 1\right\}. \tag{6.7}$$

Defining $b = [\|E_0\|_M/\omega_0, \ldots, \|E_m\|_M/\omega_m]^T$, we have $\|\triangle\mathrm{P}\| = \|b\|_v$. By the triangular inequality,

$$|y^H \triangle\mathrm{P}(\lambda)x| \leq \sum_{j=0}^{m} |\lambda|^j |y^H E_j x|. \tag{6.8}$$

With a suitable scaling of E_j by a complex number of modulus 1, we can assume without loss of generality that equality holds in (6.8). Hence,

$$\sup_{\|\triangle\mathrm{P}\|\leq 1} |y^H \triangle\mathrm{P}(\lambda)x| = \sup_{\|b\|_v \leq 1} \sum_{j=0}^{m} |\lambda|^j \sup_{\|E_j\|_M = \omega_j b_j} |y^H E_j x|. \tag{6.9}$$

Using the particular perturbation $E_j = \omega_j b_j y x^H$, it can be easily seen that the inner supremum is $\omega_j b_j$ and hence

$$\sup_{\|\triangle\mathrm{P}\|\leq 1} |y^H \triangle\mathrm{P}(\lambda)x| = \sup_{\|b\|_v \leq 1} \left|[\omega_0, \omega_1|\lambda|, \ldots, \omega_m|\lambda|^m]b\right| = \|[\omega_0,\, \omega_1|\lambda|,\, \ldots, \omega_m|\lambda|^m]\|_d,$$

which completes the proof. ■

From a practical point of view, measuring the perturbations of the individual coefficients of the polynomial separably makes a lot of sense and thus the choice $\|\cdot\|_v \equiv \|\cdot\|_\infty$ seems to be most natural. However, it turns out – especially when considering structured condition numbers – that more elegant results are obtained with the choice $\|\cdot\|_v \equiv \|\cdot\|_2$, which we will use throughout the rest of this paper. In this case, the expression (6.6) takes the form

$$\kappa_{\mathrm{P}}(\lambda) = \frac{\|[\omega_0,\, \omega_1\lambda,\, \ldots, \omega_m\lambda^m]\|_2}{|y^H \mathrm{P}'(\lambda)x|}, \tag{6.10}$$

see also [1, 8].

If $\lambda = \infty$ is a simple eigenvalue of P, a suitable condition number can be defined as

$$\kappa_{\mathrm{P}}(\infty) := \lim_{\epsilon\to 0} \sup\{1/|\hat{\lambda}\epsilon| : \ \|\triangle\mathrm{P}\| \leq \epsilon\},$$

and, following the arguments above,

$$\kappa_{\mathrm{P}}(\infty) = \omega_m/|y^H A_{m-1}x|$$

for any p-norm $\|\cdot\|_v$. Note that this discrimination between finite and infinite disappears when homogenizing P as in [26] or measuring the distance between perturbed eigenvalues with the chordal metric as in [87]. In order to keep the presentation simple, we have decided not to use these concepts.

The rest of this section is concerned with quantifying the effect on the condition number if we restrict the perturbation $\triangle\mathrm{P}$ to a subset $\mathbb{S}$ of the space of all $n \times n$ matrix polynomials

TH-789_04612301

of degree at most m.

Definition 6.2.2. *Let λ be a simple finite eigenvalue of a matrix polynomial* P *with normalized right and left eigenvectors x and y. Then the* structured condition number *of λ with respect to $\mathbb{S}$ is defined as*

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) := \lim_{\epsilon \to 0} \sup \left\{ \frac{|\hat{\lambda} - \lambda|}{\epsilon} : \triangle\mathrm{P} \in \mathbb{S}, \ \|\triangle\mathrm{P}\| \leq \epsilon \right\} \tag{6.11}$$

For infinite λ, $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) := \lim_{\epsilon \to 0} \sup\{1/|\hat{\lambda}\epsilon| : \triangle\mathrm{P} \in \mathbb{S}, \ \|\triangle\mathrm{P}\| \leq \epsilon\}$.

If $\mathbb{S}$ is star-shaped, the expansion (6.3) can be used to show

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \frac{1}{|y^H \mathrm{P}'(\lambda)x|} \sup \left\{ |y^H \triangle\mathrm{P}(\lambda)x| : \triangle\mathrm{P} \in \mathbb{S}, \ \|\triangle\mathrm{P}\| \leq 1 \right\} \tag{6.12}$$

and

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = \frac{1}{|y^H A_{m-1}x|} \sup \left\{ |y^H E_m x| : \triangle\mathrm{P} \in \mathbb{S}, \ \|E_m\|_M \leq \omega_m \right\}. \tag{6.13}$$

6.2.1 Structured first-order perturbation sets

To proceed from (6.12) we need to find the maximal absolute magnitude of elements from the set

$$\left\{ y^H \triangle\mathrm{P}(\lambda)x = y^H E_0 x + \lambda y^H E_1 x + \cdots \lambda^m y^H E_m x : \triangle\mathrm{P} \in \mathbb{S}, \ \|\triangle\mathrm{P}\| \leq 1 \right\} \tag{6.14}$$

It is therefore of interest to study the nature of the set $\{y^H E x : E \in \mathbb{E}, \ \|E\|_M \leq 1\}$ with respect to some $\mathbb{E} \subseteq \mathbb{C}^{n \times n}$. The following theorem by Karow [49] provides explicit descriptions of this set for certain $\mathbb{E}$. We use $\cong$ to denote the natural isomorphism between $\mathbb{C}$ and $\mathbb{R}^2$.

Theorem 6.2.3. *Let $\mathbb{K}(\mathbb{E}, x, y) := \{y^H E x : E \in \mathbb{E}, \ \|E\|_M \leq 1\}$ for $x, y \in \mathbb{C}^n$ with $\|x\|_2 = \|y\|_2 = 1$ and some $\mathbb{E} \subseteq \mathbb{C}^{n \times n}$. Provided that $\|\cdot\|_M \in \{\|\cdot\|_2, \|\cdot\|_F\}$, the set $\mathbb{K}(\mathbb{E}, x, y)$ is an ellipse taking the form*

$$\mathbb{K}(\mathbb{E}, x, y) \cong \mathbb{K}(\alpha, \beta) := \left\{ K(\alpha, \beta)\xi : \ \xi \in \mathbb{R}^2, \|\xi\|_2 \leq 1 \right\}, \quad K(\alpha, \beta) \in \mathbb{R}^{2 \times 2}, \tag{6.15}$$

for the cases that $\mathbb{E}$ consists of all complex ($\mathbb{E} = \mathbb{C}^{n \times n}$), real ($\mathbb{E} = \mathbb{R}^{n \times n}$), Hermitian ($\mathbb{E} = \mathsf{Herm}$), complex symmetric ($\mathbb{E} = \mathsf{symm}$), and complex skew-symmetric ($\mathbb{E} = \mathsf{skew}$), real symmetric (only for $\|\cdot\|_M \equiv \|\cdot\|_F$), and real skew-symmetric matrices. The matrix $K(\alpha, \beta)$ defining the ellipse in (6.15) can be written as

$$K(\alpha, \beta) = \begin{bmatrix} \cos\phi/2 & \sin\phi/2 \\ -\sin\phi/2 & \cos\phi/2 \end{bmatrix} \begin{bmatrix} \sqrt{\alpha + |\beta|} & 0 \\ 0 & \sqrt{\alpha - |\beta|} \end{bmatrix} \tag{6.16}$$

with some of the parameter configurations α, β listed in Table 6.1, and $\phi = \arg(\beta)$.

Note that (6.15)–(6.16) describes an ellipse with semiaxes $\sqrt{\alpha + |\beta|}$, $\sqrt{\alpha - |\beta|}$, rotated by the angle $\phi/2$. The Minkowski sum of ellipses is still convex but in general not an ellipse [56]. Finding the maximal element in (6.14) is equivalent to finding the maximal element in the Minkowski sum.

TH-789_04612301

	$\Vert\cdot\Vert_M \equiv \Vert\cdot\Vert_2$		$\Vert\cdot\Vert_M \equiv \Vert\cdot\Vert_F$	
$\mathbb{E}$	α	β	α	β
$\mathbb{C}^{n\times n}$	1	0	1	0
Herm	$1-\frac{1}{2}\lvert y^H x\rvert^2$	$\frac{1}{2}(y^H x)^2$	$\frac{1}{2}$	$\frac{1}{2}(y^H x)^2$
symm	1	0	$\frac{1}{2}(1+\lvert y^T x\rvert^2)$	0
skew	$1-\lvert y^T x\rvert^2$	0	$\frac{1}{2}(1-\lvert y^T x\rvert^2)$	0

Table 6.1: Parameters defining the ellipse (6.15)

Lemma 6.2.4. *Let $\mathbb{K}(\alpha_0,\beta_0),\ldots,\mathbb{K}(\alpha_m,\beta_m)$ be ellipses of the form (6.15)–(6.16). Define*

$$\sigma := \sup_{\substack{b_0,\ldots,b_m\in\mathbb{R}\\ b_0^2+\cdots+b_m^2\le 1}} \sup\left\{\|s\|_2 : \ s\in b_0\mathbb{K}(\alpha_0,\beta_0)+\cdots+b_m\mathbb{K}(\alpha_m,\beta_m)\right\} \tag{6.17}$$

using the Minowski sum of sets. Then

$$\sigma = \|[K(\alpha_0,\beta_0),\ldots,K(\alpha_m,\beta_m)]\|_2, \tag{6.18}$$

and

$$\sqrt{\alpha_0+\cdots+\alpha_m} \le \sigma \le \sqrt{2}\sqrt{\alpha_0+\cdots+\alpha_m}. \tag{6.19}$$

Proof: By the definition of $\mathbb{K}(\alpha_j,\beta_j)$, it follows that

$$\begin{aligned}
\sigma &= \sup_{\substack{b_j\in\mathbb{R}\\ b_0^2+\cdots+b_m^2\le 1}} \ \sup_{\substack{\xi_j\in\mathbb{R}^2\\ \|\xi_j\|_2\le 1}} \left\|b_0K(\alpha_0,\beta_0)\xi_0+\cdots+b_mK(\alpha_m,\beta_m)\xi_m\right\|_2 \\
&= \sup_{\substack{b_j\in\mathbb{R}\\ b_0^2+\cdots+b_m^2\le 1}} \ \sup_{\substack{\tilde\xi_j\in\mathbb{R}^2\\ \|\tilde\xi_j\|_2\le b_j}} \left\|K(\alpha_0,\beta_0)\tilde\xi_0+\cdots+K(\alpha_m,\beta_m)\tilde\xi_m\right\|_2 \\
&= \sup_{\substack{\tilde\xi_j\in\mathbb{R}^2\\ \|\tilde\xi_0\|_2+\cdots+\|\tilde\xi_m\|_2\le 1}} \left\|K(\alpha_0,\beta_0)\tilde\xi_0+\cdots+K(\alpha_m,\beta_m)\tilde\xi_m\right\|_2 \\
&= \left\|\left[K(\alpha_0,\beta_0),\ \ldots,\ K(\alpha_m,\beta_m)\right]\right\|_2.
\end{aligned}$$

by simply applying the definition of the matrix 2-norm. The inequality (6.19) then follows from the well-known bound

$$\frac{1}{\sqrt{2}}\|[K(\alpha_0,\beta_0),\ldots,K(\alpha_m,\beta_m)]\|_F \le \sigma \le \|[K(\alpha_0,\beta_0),\ldots,K(\alpha_m,\beta_m)]\|_F$$

and using the fact that:

$$\|[K(\alpha_0,\beta_0),\ldots,K(\alpha_m,\beta_m)]\|_F^2 = \sum_{j=0}^{m}\|K(\alpha_j,\beta_j)\|_F^2 = \sum_{j=0}^{m} 2\alpha_j.$$

■

It is instructive to rederive the expression (6.10) for the unstructured condition number from Lemma 6.2.4. Starting from Equation (6.7), we insert the definition (6.4) of $\|\triangle \mathrm{P}\|$ for

TH-789_04612301

$\|\cdot\|_M \equiv \|\cdot\|_2$, $\|\cdot\|_v \equiv \|\cdot\|_2$, and obtain

$$\begin{aligned} \sigma &= \sup\left\{|y^H \triangle \mathrm{P}(\lambda)x| : \|\triangle \mathrm{P}\| \leq 1\right\} \\ &= \sup_{\substack{b_0^2+\cdots+b_m^2\leq 1 \\ \|E_0\|^2\leq b_0,\ldots,\|E_m\|^2\leq b_m}} \left|\sum_{j=0}^{m} \omega_j \lambda^j y^H E_j x\right| \\ &= \sup_{b_0^2+\cdots+b_m^2\leq 1} \sup\left\{|s| : \; s \in \sum_{j=0}^{m} b_j \omega_j \lambda_j \mathbb{K}(\mathbb{C}^{n\times n}, x, y)\right\}. \end{aligned} \tag{6.20}$$

By Theorem 6.2.3, $\mathbb{K}(\mathbb{C}^{n\times n}, x, y) \cong \mathbb{K}(1, 0)$ and, since a disk is invariant under rotation, $\omega_j \lambda^j \mathbb{K}(\mathbb{C}^{n\times n}, x, y) \cong \mathbb{K}(\omega_j^2 |\lambda|^{2j}, 0)$. Applying Lemma 6.2.4 yields

$$\sigma = \left\|\left[K(\omega_0^2, 0), K(\omega_1^2|\lambda|^2, 0), \ldots, K(\omega_m^2|\lambda|^{2m}, 0)\right]\right\|_2 = \left\|\left[\omega_0, \omega_1\lambda, \ldots, \omega_m\lambda^m\right]\right\|_2,$$

which together with (6.7) results in the known expression (6.10) for $\kappa_{\mathrm{P}}(\lambda)$.

In the following sections, it will be shown that the expressions for structured condition numbers follow in a similar way as corollaries from Lemma 6.2.4. To keep the notation compact, we define

$$\sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \sup\left\{|y^H \triangle \mathrm{P}(\lambda)x| : \; \triangle \mathrm{P} \in \mathbb{S}, \; \|\triangle \mathrm{P}\| \leq 1\right\}.$$

for a star-shaped structure $\mathbb{S}$. By (6.12), $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda)/|y^H \mathrm{P}'(\lambda)x|$. Let us recall that the vector norm underlying the definition of $\|\triangle \mathrm{P}\|$ in (6.4), is chosen as $\|\cdot\|_v \equiv \|\cdot\|_2$ throughout the rest of this chapter.

6.2.2 T-symmetric matrix polynomials

No or only an insignificant decrease of the condition number can be expected when imposing complex symmetries on the perturbations of a matrix polynomial.

Corollary 6.2.5. *Let $\mathbb{S}$ denote the set of T-symmetric matrix polynomials. Then for a finite or infinite, simple eigenvalue λ of a matrix polynomial $\mathrm{P} \in \mathbb{S}$,*

1. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
2. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \frac{\sqrt{1+|y^T x|^2}}{\sqrt{2}} \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

Proof: Along the line of arguments leading to (6.20),

$$\sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \sup_{b_0^2+\cdots+b_m^2\leq 1} \left\{\|s\|_2 : \; s \in \sum_{j=0}^{m} b_j \omega_j \lambda^j \mathbb{K}(\mathsf{symm}, x, y)\right\}$$

for finite λ. As in the unstructured case, $\mathbb{K}(\mathsf{symm}, x, y) \cong \mathbb{K}(1, 0)$ for $\|\cdot\|_M \equiv \|\cdot\|_2$ by Theorem 6.2.3, and thus $\kappa_{\mathrm{P}}(\lambda) = \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$. For $\|\cdot\|_M \equiv \|\cdot\|_F$ we have

$$\mathbb{K}(\mathsf{symm}, x, y) \cong \mathbb{K}((1 + |y^T x|^2)/2, 0) = \frac{\sqrt{1 + |y^T x|^2}}{\sqrt{2}} \mathbb{K}(1, 0),$$

showing the second part of the statement. The proof for infinite λ is entirely analogous. ■

TH-789_04612301

6.2.3 T-even and T-odd matrix polynomials

To describe the structured condition numbers for T-even and T-odd polynomials in a convenient manner, we introduce the vector

$$\Lambda_\omega = \left[\omega_m \lambda^m,\, \omega_{m-1}\lambda^{m-1},\, \ldots,\, \omega_1\lambda,\, \omega_0\right]^T \tag{6.21}$$

along with the even coefficient projector

$$\Pi_e : \Lambda_\omega \mapsto \Pi_e(\Lambda_\omega) := \begin{cases} \left[\omega_m\lambda^m,\, 0,\, \omega_{m-2}\lambda^{m-2},\, 0,\, \ldots,\, \omega_2\lambda^2,\, 0,\, \omega_0\right]^T, & \text{if } m \text{ is even,} \\ \left[0, \omega_{m-1}\lambda^{m-1},\, 0,\, \omega_{m-3}\lambda^{m-3},\, \ldots,\, 0,\, \omega_0\right]^T, & \text{if } m \text{ is odd.} \end{cases} \tag{6.22}$$

The odd coefficient projection is defined analogously and can be written as $(1-\Pi_e)(\Lambda_\omega)$.

Lemma 6.2.6. *Let $\mathbb{S}$ denote the set of all T-even matrix polynomials. Then for a finite, simple eigenvalue λ of a matrix polynomial $\mathrm{P} \in \mathbb{S}$,*

1. $\kappa_\mathrm{P}^\mathbb{S}(\lambda) = \sqrt{1 - |y^Tx|^2 \frac{\|(1-\Pi_e)(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_\mathrm{P}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
2. $\kappa_\mathrm{P}^\mathbb{S}(\lambda) = \frac{1}{\sqrt{2}}\sqrt{1 - |y^Tx|^2 \frac{\|(1-\Pi_e)(\Lambda_\omega)\|_2^2 - \|\Pi_e(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_\mathrm{P}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

For an infinite, simple eigenvalue,

3. $\kappa_\mathrm{P}^\mathbb{S}(\infty) = \begin{cases} \kappa_\mathrm{P}(\infty), & \textit{if } m \textit{ is even,} \\ \sqrt{1-|y^Tx|^2}\,\kappa_P(\infty), & \textit{if } m \textit{ is odd,} \end{cases}$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
4. $\kappa_\mathrm{P}^\mathbb{S}(\infty) = \begin{cases} \frac{1}{\sqrt{2}}\sqrt{1+|y^Tx|^2}\kappa_\mathrm{P}(\infty), & \textit{if } m \textit{ is even,} \\ \frac{1}{\sqrt{2}}\sqrt{1-|y^Tx|^2}\kappa_\mathrm{P}(\infty), & \textit{if } m \textit{ is odd,} \end{cases}$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

Proof: By definition, the even coefficients of a T-even polynomial are symmetric while the odd coefficients are skew-symmetric. Thus, for finite λ,

$$\sigma_\mathrm{P}^\mathbb{S}(\lambda) = \sup_{b_0^2+\cdots+b_m^2 \le 1} \sup\Big\{\|s\|_2 : \ s \in \sum_{j \text{ even}} b_j\omega_j\lambda^j \mathbb{K}(\mathsf{symm}, x, y) + \sum_{j \text{ odd}} b_j\omega_j\lambda^j\mathbb{K}(\mathsf{skew}, x, y)\Big\}.$$

Applying Theorem 6.2.3 and Lemma 6.2.4 yields for $\|\cdot\|_M \equiv \|\cdot\|_2$,

$$\begin{aligned} \sigma_\mathrm{P}^\mathbb{S}(\lambda) &= \left\|\left[\Pi_e(\Lambda_\omega)^T \otimes K(1,0),\, (1-\Pi_e)(\Lambda_\omega)^T \otimes K(1-|y^Tx|^2, 0)\right]\right\|_2 \\ &= \left\|\left[\Pi_e(\Lambda_\omega)^T,\, \sqrt{1-|y^Tx|^2}(1-\Pi_e)(\Lambda_\omega)^T\right]\right\|_2 \\ &= \sqrt{\|\Lambda_\omega\|_2^2 - |y^Tx|^2\|(1-\Pi_e)(\Lambda_\omega)\|_2^2}, \end{aligned}$$

once again using the fact that a disk is invariant under rotation. Similarly, it follows for $\|\cdot\|_M \equiv \|\cdot\|_F$ that

$$\begin{aligned} \sigma_\mathrm{P}^\mathbb{S}(\lambda) &= \frac{1}{\sqrt{2}}\left\|\left[\sqrt{1+|y^Tx|^2}\Pi_e(\Lambda_\omega)^T,\, \sqrt{1-|y^Tx|^2}(1-\Pi_e)(\Lambda_\omega)^T\right]\right\|_2 \\ &= \frac{1}{\sqrt{2}}\sqrt{\|\Lambda_\omega\|_2^2 + |y^Tx|^2\big(\|\Pi_e(\Lambda_\omega)\|_2^2 - \|(1-\Pi_e)(\Lambda_\omega)\|_2^2\big)}. \end{aligned}$$

The result for infinite λ follows in an analogous manner. ■

TH-789_04612301

Remark 6.2.7. *Note that the statement of Lemma 6.2.6 does not assume that* P *itself is* T*-even. If we impose this additional condition then, for odd* m*,* P *has a* simple *infinite eigenvalue and only if also the size of* P *is odd, see, e.g., [54]. In this case, the skew-symmetry of* A_m *forces the infinite eigenvalue to be preserved under arbitrary structure-preserving perturbations. Hence,* $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = 0$*.*

Lemma 6.2.6 reveals that the structured condition number can only be significantly lower than the unstructured one if $|y^T x|$ *and* the ratio

$$\frac{\|(1-\Pi_e)(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2} = \frac{\sum\limits_{j\ \mathrm{odd}} \omega_j^2|\lambda|^{2j}}{\sum\limits_{j=0,\ldots,m} \omega_j^2|\lambda|^{2j}} = 1 - \frac{\sum\limits_{j\ \mathrm{even}} \omega_j^2|\lambda|^{2j}}{\sum\limits_{j=0,\ldots,m} \omega_j^2|\lambda|^{2j}}$$

are close to one. The most likely situation for the latter ratio to become close to one is when m is odd, ω_m does not vanish, and $|\lambda|$ is large.

Example 6.2.8 ([81])**.** *Let*

$$\mathrm{P}(\lambda) = I + \lambda 0 + \lambda^2 I + \lambda^3 \begin{bmatrix} 0 & 1-\phi & 0 \\ -1+\phi & 0 & \mathrm{i} \\ 0 & -\mathrm{i} & 0 \end{bmatrix}$$

with $0 < \phi < 1$*. This matrix polynomial has one eigenvalue* $\lambda_\infty = \infty$ *because of the highest coefficient, which is – as any odd-sized skew-symmetric matrix – singular. The following table additionally displays the eigenvalue* $\lambda_{\max}$ *of largest magnitude, the eigenvalue* $\lambda_{\min}$ *of smallest magnitude, as well as their unstructured and structured condition numbers for the set* $\mathbb{S}$ *of* T*-even matrix polynomials. We have chosen* $\omega_j = \|A_j\|_2$ *and* $\|\cdot\|_M \equiv \|\cdot\|_2$*.*

ϕ	10^0	10^{-3}	10^{-9}
$\kappa(\lambda_\infty)$	1	1.4×10^3	1.4×10^9
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_\infty)$	0	0	0
$\lvert\lambda_{\max}\rvert$	1.47	22.4	2.2×10^4
$\kappa_{\mathrm{P}}(\lambda_{\max})$	1.12	3.5×10^5	3.5×10^{17}
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{\max})$	1.12	2.5×10^4	2.5×10^{13}
$\lvert\lambda_{\min}\rvert$	0.83	0.99	1.00
$\kappa_{\mathrm{P}}(\lambda_{\min})$	0.45	5.0×10^2	5.0×10^8
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{\min})$	0.45	3.5×10^2	3.5×10^8

The entries $0 = \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_\infty) \ll \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_\infty)$ *reflect the fact that the infinite eigenvalue stays intact under structure-preserving but not under general perturbations. For the largest eigenvalues, we observe a significant difference between the structured and unstructured condition numbers as* $\phi \to 0$*. In contrast, this difference becomes negligible for the smallest eigenvalues.*

Remark 6.2.9. *For* even m*, the structured eigenvalue condition number of a* T*-even polynomial is usually close to the unstructured one. For example if all weights are equal,* $\|(1-\Pi_e)(\Lambda)\|_2^2 \le \|\Lambda\|_2^2/2$ *implying* $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) \ge \kappa_{\mathrm{P}}(\lambda)/\sqrt{2}$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$*.*

For T-odd polynomials, we obtain the following analogous of Lemma 6.2.6 by simply exchanging the roles of odd and even in the proof.

TH-789_04612301

Lemma 6.2.10. *Let $\mathbb{S}$ denote the set of all T-odd matrix polynomials. Then for a finite, simple eigenvalue λ of a matrix polynomial $\mathrm{P} \in \mathbb{S}$,*

1. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
2. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \frac{1}{\sqrt{2}}\sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_\omega)\|_2^2 - \|(1-\Pi_e)(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

For an infinite, simple eigenvalue,

3. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = \begin{cases} \kappa_{\mathrm{P}}(\infty), & \text{if } m \text{ is odd,} \\ 0, & \text{if } m \text{ is even,} \end{cases}$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
4. $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = \begin{cases} \frac{1}{\sqrt{2}}\sqrt{1 + |y^T x|^2}\kappa_{\mathrm{P}}(\infty), & \text{if } m \text{ is odd,} \\ 0, & \text{if } m \text{ is even,} \end{cases}$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

Similar to the discussion above, the only situation for which $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$ can be expected to become significantly smaller than $\kappa_{\mathrm{P}}(\lambda)$ is for $|y^T x| \approx 1$ and $\lambda \approx 0$.

6.2.4 T-palindromic and T-anti-palindromic matrix polynomials

For a T-palindromic polynomial it is sensible to require that the weights in the choice of $\|\triangle \mathrm{P}\|$, see (6.4), satisfy $\omega_j = \omega_{m-j}$. This condition is tacitly assumed throughout the entire section. The Cayley transform for polynomials introduced in [68, Sec. 2.2] defines a mapping between palindromic/anti-palindromic and odd/even polynomials. As already demonstrated in [54] for the case $m = 1$, this idea can be used to transfer the results from the previous section to the (anti-)palindromic case. For the mapping to preserve the underlying norm we have to restrict ourselves to the case $\|\cdot\|_M \equiv \|\cdot\|_F$. The corresponding coefficient projections are given by $\Pi_\pm : \Lambda_\omega \mapsto \Pi_s(\Lambda_\omega)$ with

$$\Pi_\pm(\Lambda_\omega) := \begin{cases} \left[\omega_0 \frac{\lambda^m \pm 1}{\sqrt{2}}, \ldots, \omega_{m/2-1} \frac{\lambda^{m/2+1} \pm \lambda^{m/2-1}}{\sqrt{2}}, \omega_{m/2} \frac{\lambda^{m/2} \pm \lambda^{m/2}}{2}\right]^T & \text{if } m \text{ is even,} \\ \left[\omega_0 \frac{\lambda^m \pm 1}{\sqrt{2}}, \ldots, \omega_{(m-1)/2} \frac{\lambda^{(m+1)/2} \pm \lambda^{(m-1)/2}}{\sqrt{2}}\right]^T, & \text{if } m \text{ is odd.} \end{cases} \tag{6.23}$$

Note that $\|\Pi_+(\Lambda_\omega)\|_2^2 + \|\Pi_-(\Lambda_\omega)\|_2^2 = \|\Lambda_\omega\|_2^2$.

Lemma 6.2.11. *Let $\mathbb{S}$ denote the set of all T-palindromic matrix polynomials. Then for a finite, simple eigenvalue λ of a matrix polynomial $\mathrm{P} \in \mathbb{S}$, with $\|\cdot\|_M \equiv \|\cdot\|_F$,*

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \frac{1}{\sqrt{2}}\sqrt{1 + |y^T x|^2 \frac{\|\Pi_+(\Lambda_\omega)\|_2^2 - \|\Pi_-(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_{\mathrm{P}}(\lambda).$$

For an infinite, simple eigenvalue, $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = \kappa_{\mathrm{P}}(\infty)$.

Proof: Assume m is odd. For $\triangle \mathrm{P} \in \mathbb{S}$,

$$\begin{aligned} \triangle \mathrm{P}(\lambda) &= \sum_{j=0}^{(m-1)/2} \lambda^j E_j + \sum_{j=0}^{(m-1)/2} \lambda^{m-j} E_j^T \\ &= \sum_{j=0}^{(m-1)/2} \frac{\lambda^j + \lambda^{m-j}}{\sqrt{2}} \frac{E_j + E_j^T}{\sqrt{2}} + \sum_{j=0}^{(m-1)/2} \frac{\lambda^j - \lambda^{m-j}}{\sqrt{2}} \frac{E_j - E_j^T}{\sqrt{2}}. \end{aligned}$$

TH-789_04612301

Let us introduce the auxiliary polynomial

$$\triangle\widetilde{\mathrm{P}}(\mu) = \sum_{j=0}^{(m-1)/2} \mu^{2j} S_j + \sum_{j=0}^{(m-1)/2} \mu^{2j+1} W_j, \quad S_j = \frac{E_j + E_j^T}{\sqrt{2}}, \; W_j = \frac{E_j - E_j^T}{\sqrt{2}}.$$

Then $\widetilde{\mathrm{P}} \in \widetilde{\mathbb{S}}$, where $\widetilde{\mathbb{S}}$ denotes the set of T-even polynomials. Since symmetric and skew-symmetric matrices are orthogonal to each other with respect to the matrix inner product $\langle A, B\rangle = \mathrm{trace}(B^H A)$, we have $\|A\|_F^2 + \|A^T\|_F^2 = \|(A + A^T)/\sqrt{2}\|_F^2 + \|(A - A^T)/\sqrt{2}\|_F^2$ for any $A \in \mathbb{C}^{n\times n}$ and hence $\|\triangle\mathrm{P}\| = \|\triangle\widetilde{\mathrm{P}}\|$ for $\|\cdot\|_M \equiv \|\cdot\|_F$. This allows us to write

$$\begin{aligned}
\sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda) &= \sup\left\{|y^H \triangle\mathrm{P}(\lambda)x| : \; \triangle\mathrm{P} \in \mathbb{S}, \; \|\triangle\mathrm{P}\| \le 1\right\} \\
&= \sup\left\{\left|\sum \frac{\lambda^j + \lambda^{m-j}}{\sqrt{2}} y^H S_j x + \sum \frac{\lambda^j - \lambda^{m-j}}{\sqrt{2}} y^H W_j x\right| : \; \triangle\widetilde{\mathrm{P}} \in \widetilde{\mathbb{S}}, \; \|\triangle\widetilde{\mathrm{P}}\| \le 1\right\} \\
&= \frac{1}{\sqrt{2}} \sup_{b_0^2+\cdots+b_m^2 \le 1} \Big\{\|s\|_2 : \; s \in \sum b_j \omega_j (\lambda^j + \lambda^{m-j}) \mathbb{K}(\mathsf{symm}, x, y) \\
&\qquad\qquad + \sum b_{(m-1)/2+j}\omega_j(\lambda^j - \lambda^{m-j})\mathbb{K}(\mathsf{skew}, x, y)\Big\} \\
&= \frac{1}{2}\sqrt{(1 + |y^T x|^2)\sum \omega_j^2 |\lambda^j + \lambda^{m-j}|^2 + (1 - |y^T x|^2)\sum \omega_j^2 |\lambda^j - \lambda^{m-j}|^2} \\
&= \frac{1}{\sqrt{2}}\sqrt{(1 + |y^T x|^2)\|\Pi_+(\Lambda_\omega)\|_2^2 + (1 - |y^T x|^2)\|\Pi_+(\Lambda_\omega)\|_2^2} \\
&= \frac{1}{\sqrt{2}}\sqrt{\|\Lambda_\omega\|_2^2 + |y^T x|^2(\|\Pi_+(\Lambda_\omega)\|_2^2 - \|\Pi_-(\Lambda_\omega)\|_2^2)},
\end{aligned}$$

where we used Theorem 6.2.3 and Lemma 6.2.4.

For even m the proof is almost the same; with the only difference that the transformation leaves the complex symmetric middle coefficient $A_{m/2}$ unaltered.

For $\lambda = \infty$, observe that the corresponding optimization problem (6.13) involves only a single coefficient of the polynomial and hence palindromic structure has no effect on the condition number. ■

From the result of Lemma 6.2.11 it follows that a large difference between the structured and unstructured condition numbers for T-palindromic matrix polynomials may occur when $|y^T x|$ is close to one, *and* $\|\Pi_+(\Lambda_\omega)\|_2$ is close to zero. Assuming that all weights are positive, the latter condition implies that m is odd and and $\lambda \approx -1$. An instance of such a case is given by a variation of Example 6.2.8.

Example 6.2.12. *Consider the T-palindromic matrix polynomial*

$$\mathrm{P}(\lambda) = \begin{bmatrix} 1 & 1-\phi & 0 \\ -1+\phi & 1 & \mathrm{i} \\ 0 & -\mathrm{i} & 1 \end{bmatrix} + \lambda I + \lambda^2 I - \lambda^3 \begin{bmatrix} 1 & 1-\phi & 0 \\ -1+\phi & 1 & \mathrm{i} \\ 0 & -\mathrm{i} & 1 \end{bmatrix}$$

with $0 < \phi < 1$. An odd-sized T-palindromic matrix polynomial, P *has the eigenvalue $\lambda_{-1} = -1$. The following table additionally displays one eigenvalue λ_{close} closest to -1, an eigenvalue $\lambda_{\min}$ of smallest magnitude, as well as their unstructured and structured condition numbers for the set $\mathbb{S}$ of T-palindromic matrix polynomials. We have chosen $\omega_j = \|A_j\|_F$ and $\|\cdot\|_M \equiv \|\cdot\|_F$.*

TH-789_04612301

ϕ	10^{-1}	10^{-4}	10^{-8}
$\kappa(\lambda_{-1})$	20.9	2.2×10^4	2.2×10^8
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{-1})$	0	0	0
$\|1+\lambda_{\mathrm{close}}\|$	0.39	1.4×10^{-2}	1.4×10^{-4}
$\kappa_{\mathrm{P}}(\lambda_{\mathrm{close}})$	11.1	1.1×10^4	1.1×10^8
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{\mathrm{closer}})$	6.38	2.5×10^2	2.6×10^4
$\|1-\lambda_{\mathrm{min}}\|$	1.25	1.41	1.41
$\kappa_{\mathrm{P}}(\lambda_{\mathrm{min}})$	7.92	7.9×10^3	7.9×10^7
$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{\mathrm{min}})$	5.75	5.6×10^3	5.6×10^7

The entries $0 = \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda_{-1}) \ll \kappa_{\mathrm{P}}(\lambda_{-1})$ reflect the fact that the eigenvalue -1 remains intact under structure-preserving but not under general perturbations. Also, eigenvalues close to -1 benefit from a significantly lower structured condition numbers as $\phi \to 0$. In contrast, only a practically irrelevant benefit is revealed for the eigenvalue λ_{min} not close to -1.

Structured eigenvalue condition numbers for T-anti-palindromic matrix polynomials can be derived in the same way as in Lemma 6.2.11.

Lemma 6.2.13. *Let $\mathbb{S}$ denote the set of all T-anti-palindromic matrix polynomials. Then for a finite, simple eigenvalue λ of a matrix polynomial $\mathrm{P} \in \mathbb{S}$, with $\|\cdot\|_M \equiv \|\cdot\|_F$,*

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \frac{1}{\sqrt{2}}\sqrt{1 - |y^T x|^2 \frac{\|\Pi_+(\Lambda_\omega)\|_2^2 - \|\Pi_-(\Lambda_\omega)\|_2^2}{\|\Lambda_\omega\|_2^2}}\, \kappa_{\mathrm{P}}(\lambda).$$

For an infinite, simple eigenvalue, $\kappa_{\mathrm{P}}^{\mathbb{S}}(\infty) = \kappa_{\mathrm{P}}(\infty)$.

6.2.5 H-Hermitian matrix polynomials

The derivations in the previous sections were greatly simplified by the fact that the first-order perturbation sets under consideration were disks. For the set of Hermitian perturbations, however, $y^H E_j x$ forms truly an ellipse. Still, a computable expression is provided by (6.18) from Lemma 6.2.4. However, the explicit formulas derived from this expression take a very technical form and provide little immediate intuition on the difference between the structured and unstructured condition number. Therefore, we will work with the bound (6.19) instead.

Lemma 6.2.14. *Let $\mathbb{S}$ denote the set of all H-Hermitian matrix polynomials. Then for a finite or infinite, simple eigenvalue of a matrix polynomial $\mathrm{P} \in \mathbb{S}$,*

1. $\sqrt{1 - \frac{1}{2}|y^H x|^2}\, \kappa_{\mathrm{P}}(\lambda) \le \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) \le \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_2$, *and*
2. $\kappa_{\mathrm{P}}(\lambda)/\sqrt{2} \le \kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) \le \kappa_{\mathrm{P}}(\lambda)$ *for* $\|\cdot\|_M \equiv \|\cdot\|_F$.

Proof: Let $\|\cdot\|_M \equiv \|\cdot\|_F$. Then Theorem 6.2.3 states

$$\mathbb{K}(\mathsf{Herm}, x, y) \cong \mathbb{K}(1/2, (y^H x)^2/2).$$

Consequently,

$$\omega_j \lambda^j \mathbb{K}(\mathsf{Herm}, x, y) \cong \mathbb{K}(\omega_j^2 |\lambda|^{2j}/2, \omega_j^2 \lambda^{2j} (y^H x)^2/2),$$

TH-789_04612301

which implies

$$\sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \sup_{b_0^2+\cdots+b_m^2 \le 1} \Big\{ \|s\|_2 : \ s \in \sum_{j=0}^{m} b_j \mathbb{K}(\omega_j^2 \lambda^{2j}/2, \omega_j^2 \lambda^{2j} (y^H x)^2/2) \Big\}.$$

By Lemma 6.2.4,

$$\frac{1}{\sqrt{2}} \|\Lambda_\omega\|_2 \le \sigma_{\mathrm{P}}^{\mathbb{S}}(\lambda) \le \|\Lambda_\omega\|_2.$$

The proof for the case $\|\cdot\|_M \equiv \|\cdot\|_2$ is analogous. ■

Remark 6.2.15. *Since H-Hermitian and H-skew-Hermitian matrices are related by multiplication with* i*, which simply rotates the first-order perturbation set by* 90 *degrees, a slight modification of the proof shows that the statement of Lemma 6.2.14 remains true when $\mathbb{S}$ denotes the space of H-odd or H-even polynomials. This can in turn be used – as in the proof of Lemma 6.2.11 – to show that also for H-(anti-)palindromic polynomials there is at most an insignificant difference between the structured and unstructured eigenvalue condition numbers.*

6.3 Condition numbers for linearizations

As already mentioned in the introduction, polynomial eigenvalue problems are often solved by first linearizing the matrix polynomial into a larger matrix pencil. Of the classes of linearizations proposed in the literature, the vector spaces $\mathbb{DL}(\mathrm{P})$, defined in (1.4) are particularly amenable to further analysis, while offering a degree of generality that is often sufficient in applications.

Definition 6.3.1. *Let $\Lambda_{m-1} = [\lambda^{m-1}, \lambda^{m-2} \dots \lambda, 1]^T$ and let* P *be a matrix polynomial of degree m. Then a matrix pencil $\mathrm{L}(\lambda) = \lambda X + Y \in \mathbb{C}^{mn\times mn}$ is in $\mathbb{DL}(\mathrm{P})$ if there is a so called* ansatz vector *$v \in \mathbb{C}^m$ satisfying*

$$\mathrm{L}(\lambda) \cdot (\Lambda_{m-1} \otimes I) = v \otimes P(\lambda) \quad \textit{and} \quad (\Lambda_{m-1}^T \otimes I) \cdot \mathrm{L}(\lambda) = v^T \otimes P(\lambda).$$

It is easy to see that the ansatz vector v is uniquely determined by $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. As stated earlier, it is shown in[67, Thm. 6.7] that $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ is a linearization of P if and only if none of the eigenvalues of P is a root of the v-polynomial $\mathsf{p}(\lambda; v)$ defined in (1.5) associated with the ansatz vector $v = [v_1, \dots, v_m]^T$. If P has eigenvalue ∞, this condition should be read as $v_1 \ne 0$. Apart from this elegant characterization, probably the most important property of $\mathbb{DL}(\mathrm{P})$ is that it leads to a simple one-to-one relation between the eigenvectors of P and $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. To keep the notation compact, we define Λ_{m-1} as in Definition 6.3.1 for finite λ and set $\Lambda_{m-1} = [1, 0, \dots, 0]^T$ for $\lambda = \infty$.

By Theorem 1.2.13 we know that $x \ne 0$ is a right eigenvector of P associated with an eigenvalue λ if and only if $\Lambda_{m-1} \otimes x$ is a right eigenvector of L associated with λ. Similarly, $y \ne 0$ is a left eigenvector of P associated with an eigenvalue λ if and only if $\overline{\Lambda}_{m-1} \otimes y$ is a left eigenvector of L associated with λ. As a matrix pencil $\mathrm{L}(\lambda) = \lambda X + Y$ is a special case of a matrix polynomial, we can use the results of Section 6.2 to study the (structured) eigenvalue condition numbers of L. In the unstructured case, Lemma 6.2.1 together with Theorem 1.2.13 imply for any unitarily invariant norm $\|\cdot\|_M$ and $\|\cdot\|_v \equiv \|\cdot\|_2$ the following formula.

TH-789_04612301

Lemma 6.3.2. *Let λ be a finite, simple eigenvalue of a matrix polynomial P with normalized right and left eigenvectors x and y. Then the eigenvalue condition number $\kappa_{\mathrm{L}}(\lambda)$ for a linearization $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ with ansatz vector v satisfies*

$$\kappa_{\mathrm{L}}(\lambda) = \frac{\sqrt{1+|\lambda|^2}}{|\mathsf{p}(\lambda;v)|} \cdot \frac{\|\Lambda_{m-1}\|_2^2}{|y^H P'(\lambda)x|} = \frac{\|(1,\,\lambda)\|_2 \|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|\,\|\Lambda_m\|_2}\kappa_{\mathrm{P}}(\lambda),$$

provided that the perturbations $\triangle\mathrm{L} = \triangle X + \lambda\triangle Y$ are measured in the norm $\|\triangle\mathrm{L}\| = \sqrt{\|\triangle X\|_M^2 + \|\triangle Y\|_M^2}$ for a unitarily invariant norm $\|\cdot\|_M$. Thus, we have

$$\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda;v)|} \leq \frac{\kappa_{\mathrm{L}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda;v)|}.$$

Proof: A similar formula for the case $\|\cdot\|_v \equiv \|\cdot\|_1$ can be found in [41, Section 3]. The proof for the case $\|\cdot\|_v \equiv \|\cdot\|_2$ is almost identical and therefore omitted.

It is straightforward to see

$$1 \leq \frac{\sqrt{1+|\lambda|^2}\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2} \leq \sqrt{2}. \tag{6.24}$$

Hence the result follows. ■

Remark 6.3.3. *To simplify the analysis, we will assume that the weights $\omega_0, \ldots, \omega_m$ in the definition of $\|\triangle\mathrm{P}\|$ are all equal to 1 for the rest of this chapter. This assumption is only justified if* P *is not badly scaled, i.e., the norms of the coefficients of* P *do not vary significantly. To a certain extent, bad scaling can be overcome by rescaling the matrix polynomial before linearization, see [29, 39, 41, 42].*

Given an ansatz vector v, for the rest of the chapter, we set

$$\delta_v := \frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda;v)|}. \tag{6.25}$$

Obviously $\delta_v \geq 1$. Note that $\mathsf{p}(\lambda;v) \neq 0$ and hence $\delta_v < \infty$ when L is the linearization of P corresponding to v. Lemma 6.3.2 shows that linearizing P by the pencil $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ corresponding to an ansatz vector v invariably increases the condition number of a simple eigenvalue of P at least by a factor of δ_v and at most by a factor of $\sqrt{2}\delta_v$. Thus δ_v serves as a growth factor of condition number of a simple eigenvalue of P when P is linearized by the pencil L associated with v.

Since $\mathsf{p}(\lambda;v) = \Lambda_{m-1}^T v$, it follows from the Cauchy-Schwartz inequality that among all ansatz vectors with $\|v\|_2 = 1$ the vector $v = \overline{\Lambda}_{m-1}/\|\Lambda_{m-1}\|_2$ minimizes δ_v and, hence, for this particular choice of v we have $\delta_v = 1$ and

$$\kappa_{\mathrm{P}}(\lambda) \leq \kappa_{\mathrm{L}}(\lambda) \leq \sqrt{2}\,\kappa_{\mathrm{P}}(\lambda).$$

Let us emphasize that this result is primarily of theoretical interest as the optimal choice of v depends on the (typically unknown) eigenvalue λ. A practically more useful recipe is to

TH-789_04612301

choose $v = [1, 0, \ldots, 0]^T$ if $|\lambda| \geq 1$ and $v = [0, \ldots, 0, 1]^T$ if $|\lambda| \leq 1$. In both cases,

$$\delta_v = \frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; \overline{\Lambda}_{m-1}/\|\Lambda_{m-1}\|_2)|} \leq \sqrt{m} \tag{6.26}$$

and therefore $\kappa_{\mathrm{P}}(\lambda) \leq \kappa_{\mathrm{L}}(\lambda) \leq \sqrt{2m}\, \kappa_{\mathrm{P}}(\lambda)$.

In the following section, the discussion above shall be extended to structured linearizations and condition numbers.

6.4 Structured condition numbers for linearizations

If the polynomial P is structured, its linearization $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ should reflect this structure. Table 6.3 summarizes existing results on the conditions the ansatz vector v should satisfy for this purpose. These conditions can be found in [40, Thm 3.4] for symmetric polynomials, in [40, Thm. 6.1] for Hermitian polynomials. If, for example, a structure-preserving method

structure of P	structure of L	ansatz vector
T-symmetric	T-symmetric	$v \in \mathbb{C}^m$
H-Hermitian	H-Hermitian	$v \in \mathbb{R}^m$

Table 6.2: Conditions the ansatz vector v needs to satisfy in order to yield a structured linearization $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ for a structured polynomial P.

is used for computing the eigenvalues of a structured linearization L then the structured condition number of L is an appropriate measure for the influence of roundoff error on the accuracy of the computed eigenvalues. It is therefore of interest to choose L such that the structured condition number is minimized.

Observe that if L is the structured linearization of P corresponding to an ansatz vector v then by Lemma 6.3.2, we have

$$\frac{\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \frac{\kappa_{\mathrm{L}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\, \delta_v$$

for both $\|\cdot\|_M \equiv \|\cdot\|_F$ and $\|\cdot\|_M \equiv \|\cdot\|_2$.

6.4.1 T-symmetric matrix polynomials

For a T-symmetric matrix polynomial P, any ansatz vector v yields a T-symmetric linearization. Thus, we are free to use the optimal choice $v = \overline{\Lambda}_{m-1}/\|\Lambda_{m-1}\|_2$ from Section 6.3. Combined with Corollary 6.2.5, which states that there is (almost) no difference between structured and unstructured condition numbers, we have the following result.

Theorem 6.4.1. *Let $\mathbb{S}$ denote the set of T-symmetric matrix polynomials. Let λ be a finite or infinite, simple eigenvalue of a matrix polynomial $\mathrm{P} \in \mathbb{S}$. Then for the linearization $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ corresponding to an ansatz vector v, we have*

$$\delta_v \leq \frac{\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)} \leq \sqrt{2}\, \delta_v$$

TH-789_04612301

for $\|\cdot\|_M \equiv \|\cdot\|_2$ *and* $\|\cdot\|_M \equiv \|\cdot\|_F$*. In particular, for* $v = \overline{\Lambda}_{m-1}/\|\Lambda_{m-1}\|_2$*, we have*

$$\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) \leq \kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda) \leq \sqrt{2}\,\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda). \tag{6.27}$$

Proof: For $\|\cdot\|_M \equiv \|\cdot\|_2$, we have $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \kappa_{\mathrm{P}}(\lambda)$ and $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda) = \kappa_{\mathrm{L}}(\lambda)$. Hence the result follows from Lemma 6.3.2. For $\|\cdot\|_M \equiv \|\cdot\|_F$, the additional factors appearing in Corollary 6.2.5 are the same for $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$ and $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)$. This can be seen as follows. According to Theorem 1.2.13, the normalized right and left eigenvectors of the linearization take the form $\tilde{x} = \Lambda_{m-1} \otimes x/\|\Lambda_{m-1}\|_2$, $\tilde{y} = \overline{\Lambda_{m-1}} \otimes y/\|\Lambda_{m-1}\|_2$. Thus,

$$\tilde{y}^T\tilde{x} = \frac{\overline{\Lambda_{m-1}}^T\Lambda_{m-1}}{\|\Lambda_{m-1}\|_2^2} y^T x = y^T x, \tag{6.28}$$

concluding the proof. ■

Assuming that $\|\cdot\|_M \equiv \|\cdot\|_F$ and $\|\cdot\|_M \equiv \|\cdot\|_2$ the results discussed above present growth of structured condition number of a simple eigenvalue of $\mathrm{P} \in \mathbb{S}$ when P is linearized by a structured pencil in $\mathbb{DL}(\mathrm{P})$. However, comparing $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)$ with $\kappa_{\mathrm{P}}(\lambda)$ we have the following results for $\|\cdot\|_M \equiv \|\cdot\|_F$ and $\|\cdot\|_M \equiv \|\cdot\|_2$.

Theorem 6.4.2. *Let* $\mathbb{S}$ *denote the set of* T*-symmetric matrix polynomials. Let* λ *be a finite or infinite, simple eigenvalue of a matrix polynomial* $\mathrm{P} \in \mathbb{S}$*. Then for the linearization* $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ *corresponding to an ansatz vector* v*, we have*

1. $\|\cdot\|_M \equiv \|\cdot\|_F : \dfrac{\sqrt{1+|x^Ty|^2}}{\sqrt{2}}\delta_v \leq \dfrac{\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{1+|x^Ty|^2}\delta_v$
2. $\|\cdot\|_M \equiv \|\cdot\|_2 : \delta_v \leq \dfrac{\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_v$

Proof: First we consider $\|\cdot\|_M \equiv \|\cdot\|_F$. By Theorem 6.2.5 we have $\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda) = \dfrac{\sqrt{1+|y^Tx|^2}}{\sqrt{2}}\kappa_{\mathrm{P}}(\lambda)$. If $\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ is a T-symmetric linearization of $\mathrm{P} \in \mathbb{S}$ corresponding to the ansatz vector v then by (6.28) we have

$$\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda) = \frac{\sqrt{1+|y^Tx|^2}}{\sqrt{2}} \frac{\|\Lambda_{m-1}\|_2^2\|(\lambda,1)\|_2}{|\Lambda_{m-1}^Tv|\,\|\Lambda_m\|_2}\kappa_{\mathrm{P}}(\lambda)$$

for $\|\cdot\|_M \equiv \|\cdot\|_F$. Hence by (6.24) the desired result follows.

Next consider $\|\cdot\|_M \equiv \|\cdot\|_2$. By Theorem 6.2.5 we have

$$\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda) = \frac{\|(1,\lambda)\|_2\|\Lambda_{m-1}\|_2^2}{|\Lambda_{m-1}^Tv|\|\Lambda_m\|_2}\kappa_{\mathrm{P}}(\lambda).$$

The desired result follows by (6.24).■

6.4.2 T-even and T-odd matrix polynomials

In contrast to T-symmetric polynomials, structure-preserving linearizations for T-even and T-odd polynomials put a restriction on the choice of the ansatz vector. Conditions for structured linearization for T-even/T-odd matrix polynomial are given in Table 6.3 and can be found in [68, Table 3.2]. The conditions in second and third columns of Table 6.3 are equivalent.

TH-789_04612301

structure of P	(i) structure of $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$	(ii) $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ and ansatz vector
$*$-even	$*$-even	$\Sigma v = (v^*)^T$
	$*$-odd	$\Sigma v = -(v^*)^T$
$*$-odd	$*$-even	$\Sigma v = -(v^*)^T$
	$*$-odd	$\Sigma v = (v^*)^T$

Table 6.3: Conditions the ansatz vector v needs to satisfy to yield a structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ such that $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$, where $\Sigma \in \mathbb{R}^{m\times m}$ is defined as $\Sigma = \mathrm{diag}\{(-1)^{m-1}, (-1)^{m-2}, \ldots, (-1)^0\}$ and $* \in \{T, H\}$.

Note that $\mathrm{P} \in \mathbb{S}$ has a structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ such that the (unstructured) linearization $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. Therefore if $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ are the right and left eigenvectors of P corresponding to the eigenvalue λ then $\Lambda_{m-1} \otimes x$ and $\overline{\Lambda_{m-1}} \otimes y$ are right and left eigenvectors corresponding to the eigenvalue λ of $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$, respectively. This implies $\widetilde{x} = \Lambda_{m-1} \otimes x$ and $\widetilde{y} = \Sigma\overline{\Lambda_{m-1}} \otimes y$ are right and left eigenvectors of the structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ corresponding to the eigenvalue λ, respectively.

Since $\mathrm{L}(\lambda) \in \mathbb{L}_1(\mathrm{P})$, we have

$$\mathrm{L}(\lambda)(\Lambda_{m-1} \otimes I_n) = v \otimes \mathrm{P}(\lambda) \Rightarrow \mathrm{L}'(\lambda)(\Lambda_{m-1} \otimes I_n) + \mathrm{L}(\lambda)(\Lambda'_{m-1} \otimes I_n) = v \otimes \mathrm{P}'(\lambda). \quad (6.29)$$

This yields

$$(\Sigma\overline{\Lambda_{m-1}} \otimes y)^H \mathrm{L}'(\lambda)(\Lambda_{m-1} \otimes x) = \Lambda_{m-1}^T \Sigma v \otimes y^H \mathrm{P}'(\lambda)x \quad (6.30)$$

by multiplying $(\Sigma\overline{\Lambda_{m-1}} \otimes y)^H$ to left and $(1 \otimes x)$ to the right of (6.29).

First we derive the unstructured condition number of a simple eigenvalue λ of the structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ for a given $\mathrm{P} \in \mathbb{S}$.

Lemma 6.4.3. *Let λ be a finite, simple eigenvalue of a matrix polynomial P with normalized right and left eigenvectors x and y. Then the eigenvalue condition number $\kappa_{\mathrm{L}}(\lambda)$ for a linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ with ansatz vector v such that $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ satisfies*

$$\kappa_{\mathrm{L}}(\lambda) = \frac{\sqrt{1+|\lambda|^2}}{|\mathsf{p}(\lambda; \Sigma v)|} \cdot \frac{\|\Lambda_{m-1}\|_2^2}{|y^H P'(\lambda)x|} = \frac{\|(1,\, \lambda)\|_2 \|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda; \Sigma v)|\; \|\Lambda_m\|_2} \kappa_{\mathrm{P}}(\lambda),$$

provided that the perturbations $\triangle \mathrm{L} = \triangle X + \lambda \triangle Y$ are measured in the norm $\|\triangle \mathrm{L}\| = \sqrt{\|\triangle X\|_M^2 + \|\triangle Y\|_M^2}$ for a unitarily invariant norm $\|\cdot\|_M$. Thus, we have

$$\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; \Sigma v)|} \leq \frac{\kappa_{\mathrm{L}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; \Sigma v)|}.$$

Proof: The proof follows from Lemma 6.3.2 and (6.30).■

Note that $\delta_{\Sigma v} = \dfrac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; \Sigma v)|}$ serves as a growth factor for the unstructured condition number for a structured linearization $\mathrm{L}(\lambda) \in \mathbb{L}_1(\mathrm{P})$ such that $(\Sigma \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. Obviously $\delta_{\Sigma v} \geq 1$. Note that $\mathsf{p}(\lambda; \Sigma v) \neq 0$ and hence $\delta_{\Sigma v} < \infty$ when L is the structured linearization of P corresponding to v.

Obtaining an optimally conditioned linearization requires finding the maximum of $|\mathsf{p}(\lambda; \Sigma v)| = |\Lambda_{m-1}^T \Sigma v| =$ among all such v with $\|v\|_2 \leq 1$. This maximization problem can be addressed

TH-789_04612301

by the following basic linear algebra result.

Proposition 6.4.4. *Let $\Pi_{\mathcal{V}}$ be an orthogonal projector onto a linear subspace $\mathcal{V}$ of $\mathbb{F}^m$ with $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$. Then for $A \in \mathbb{F}^{l\times m}$,*

$$\max_{\substack{v\in\mathcal{V}\\ \|v\|_2\leq 1}} \|Av\|_2 = \|A\Pi_{\mathcal{V}}\|_2.$$

For a T-even linearization we have $\mathcal{V} = \{v \in \mathbb{C}^m : \Sigma v = v\}$ and the orthogonal projector onto $\mathcal{V}$ is given by the even coefficient projector Π_e defined in (6.22). Hence, by Proposition 6.4.4,

$$\max_{\substack{v=\Sigma v\\ \|v\|_2\leq 1}} |\mathsf{p}(\lambda;\Sigma v)| = \max_{\substack{v=\Sigma v\\ \|v\|_2\leq 1}} |\Lambda_{m-1}^T \Sigma v| = \|\Pi_e(\Lambda_{m-1})\|_2$$

where the maximum is attained by $v = \Pi_e(\overline{\Lambda}_{m-1})/\|\Pi_e(\Lambda_{m-1})\|_2$. Similarly, for a T-odd linearization,

$$\max_{\substack{v=-\Sigma v\\ \|v\|_2\leq 1}} |\mathsf{p}(\lambda;\Sigma v)| = \|(1-\Pi_e)(\Lambda_{m-1})\|_2$$

with the maximum attained by $v = (1-\Pi_e)(\overline{\Lambda}_{m-1})/\|(1-\Pi_e)(\Lambda_{m-1})\|_2$.

Also note that in contrast to the equality given in (6.28) we have the following inequality

$$\widetilde{y}^T\widetilde{x} = \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|}{\|\Lambda_{m-1}\|_2^2} y^T x \leq y^T x. \tag{6.31}$$

Theorem 6.4.5. *Let $\mathbb{S}_e$ and $\mathbb{S}_o$ denote the sets of T-even and T-odd polynomials, respectively. Let λ be a finite or infinite, simple eigenvalue of a T-even matrix polynomial P. Consider the T-even (resp. T-odd) linearizations L_e (resp. L_o) from $\mathbb{L}_1(\mathrm{P})$ corresponding to the ansatz vector $\Sigma v = v$ (resp. $\Sigma v = -v$). Then the following statements hold for $\|\cdot\|_M \equiv \|\cdot\|_2$.*

1. *If m is odd and $|\lambda| \leq 1$: $\delta_{\Sigma v} \leq \dfrac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq 2\delta_{\Sigma v}$.*
2. *If m is even and $|\lambda| \leq 1$: $\dfrac{\delta_{\Sigma v}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq 2\delta_{\Sigma v}$.*
3. *If m is even and $|\lambda| \geq 1$: $\dfrac{\delta_{\Sigma v}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq 2\delta_{\Sigma v}$.*
4. *If m is odd and $|\lambda| \geq 1$: $\delta_{\Sigma v} \leq \dfrac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq \sqrt{2}\alpha\delta_{\Sigma v}$, where $\alpha = 1$ if $y^T x = 0$ and $\alpha \leq 1/\sqrt{1-|y^T x|^2}$ if $y^T x \neq 0$.*

In particular, considering the linearizations L_e and L_o corresponding to the ansatz vectors $v = \Pi_e(\overline{\Lambda}_{m-1})/\|\Pi_e(\Lambda_{m-1})\|_2$ and $v = (1-\Pi_e)(\overline{\Lambda}_{m-1})/\|(1-\Pi_e)(\Lambda_{m-1})\|_2$, respectively, we have $\delta_{\Sigma v} \leq \sqrt{2}$.

Proof: The proof makes use of the basic relations

$$\frac{|\lambda|^2}{1+|\lambda|^2} \geq \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}, \quad \text{with equality for odd } m, \tag{6.32}$$

and

$$\|\Pi_e(\Lambda_{m-1})\| \leq \|\Lambda_{m-1}\|_2 \leq \sqrt{2}\|\Pi_e(\Lambda_{m-1})\|, \tag{6.33}$$

TH-789_04612301

which holds if either m is odd or m is even and $|\lambda| \leq 1$.

1. If m is odd, (6.32) implies – together with Lemma 6.2.6 and (6.31) – the inequality

$$\frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} = \frac{\sqrt{1 - |y^T x| \frac{|\lambda|^2}{1+|\lambda|^2} \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{1 - |y^T x| \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \geq \frac{\sqrt{1 - |y^T x| \frac{|\lambda|^2}{1+|\lambda|^2}}}{\sqrt{1 - |y^T x| \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

   Further note that $\sqrt{1 - |y^T x| \frac{|\lambda|^2}{1+|\lambda|^2} \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}} \leq 1$. Then by (6.32) we have

$$\frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq \frac{1}{\sqrt{1 - |y^T x|^2 \frac{|\lambda|^2}{1+|\lambda|^2}}} \frac{\kappa_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

   Now if $|\lambda| \leq 1$ we have $\sqrt{1 - |y^T x|^2 \frac{|\lambda|^2}{1+|\lambda|^2}} \geq \frac{1}{\sqrt{1+|\lambda|^2}} \geq \frac{1}{\sqrt{2}}$ since $|y^T x| \leq 1$. Now the desired result follows from Lemma 6.4.3.

2. If m is even and $|\lambda| \leq 1$ then (6.32) yields

$$\frac{1}{\sqrt{2}} \cdot \frac{\kappa_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)}.$$

   Now

$$\frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq \frac{1}{\sqrt{1 - |y^T x|^2 \frac{\|(I-\Pi_e(\Lambda_m))\|_2^2}{\|\Lambda_m\|_2^2}}} \frac{\kappa_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

   Note that $\sqrt{1 - |y^T x|^2 \frac{\|(I-\Pi_e(\Lambda_m))\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1 - \frac{\|\Lambda_m\|_2^2 - \|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\|\Pi_e(\Lambda_m)\|_2}{\|\Lambda_m\|_2} \geq \frac{1}{\sqrt{2}}$ since $|y^T x| \leq 1$.

   The desired result follows from Lemma 6.4.3.

3. If $|\lambda| \geq 1$ and a T-odd linearization is used then Lemma 6.2.10 yields

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} = \frac{\sqrt{1 - |y^T x|^2 \frac{1}{1+|\lambda|^2} \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{1 - |y^T x|^2 \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

   Hence

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \geq \frac{\sqrt{1 - |y^T x|^2 \frac{1}{1+|\lambda|^2}}}{\sqrt{1 - |y^T x|^2 \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

   The relation

$$\frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \leq \frac{1}{1+|\lambda|^2} \leq \frac{1}{2}$$

TH-789_04612301

holds true for even m. Consequently we have

$$\frac{1}{\sqrt{2}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)}.$$

for even m. Further

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq \frac{1}{\sqrt{1 - |y^T x|^2 \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Note that $\sqrt{1 - |y^T x|^2 \frac{\|(I-\Pi_e(\Lambda_m))\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1 - \frac{\|(I-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} \geq \frac{1}{\sqrt{2}}$ since $|y^T x| \leq 1$ and $\frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \leq \frac{1}{1+|\lambda|^2} \leq \frac{1}{2}$. Thus by Lemma 6.4.3. we obtain the desired result.

4. Next assume that m is odd and $|\lambda| \geq 1$. Then

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \geq \frac{\sqrt{1 - |y^T x|^2 \frac{|\lambda|^2}{1+|\lambda|^2}}}{\sqrt{1 - |y^T x|^2 \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}, \text{ since } -1 \geq -|\lambda|.$$

Further

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)} \leq \frac{1}{\sqrt{1 - |y^T x|^2 \frac{\|(1-\Pi_e)(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Hence the result follows from Lemma 6.4.3.

Finally, by (6.33) we have $\delta_{\Sigma v} \leq \sqrt{2}$. ■

The morale of Theorem 6.4.5 is quickly told: There is always a "good" T-even linearization (in the sense that the linearization increases the structured condition number only by a factor of $\delta_{\Sigma v}$) if either m is odd or m is even and $|\lambda| \leq 1$. In the exceptional case, when m is even and $|\lambda| \geq 1$, there is always a "good" T-odd linearization. Intuitively, the necessity of such an exceptional case becomes clear from the fact that there exists no T-even linearization for a T-even polynomial with even m and infinite eigenvalue. Even though there is a T-even linearization for even m and large but finite λ, it is not advisable to use it for numerical computations.

In practice, one does not know λ in advance and hence the linearizations used in Theorem 6.4.5 for which $\delta_{\Sigma v} \leq \sqrt{2}$ are mainly of theoretical interest. Table 6.4 provides practically more feasible recommendations on the choice of v, such that there is still at worst a slight increase of the structured condition number. The bounds in this table follow from the proof of Theorem 6.4.5 combined with (6.26). The example linearizations are taken from [68, Tables 3.4–3.6].

Theorem 6.4.6. *Let $\mathbb{S}_e$ and $\mathbb{S}_o$ denote the sets of T-even and T-odd polynomials, respectively. Let λ be a finite or infinite, simple eigenvalue of a T-odd matrix polynomial* P. *Consider the T-odd (resp. T-even) linearizations L_o (resp. L_e) from $\mathbb{L}_1(\mathrm{P})$ corresponding to the ansatz vector $\Sigma v = v$ (resp. $\Sigma v = -v$). Then the following statements hold for $\|\cdot\|_M \equiv \|\cdot\|_2$.*

1. *If m is odd and $|\lambda| \leq 1$:* $\delta_{\Sigma v} \leq \frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \sqrt{2}\sqrt{m+1}\,\delta_{\Sigma v}$.

TH-789_04612301

m	λ of interest	v	Bound on struct. cond. of linearization	Example
odd or even	$\lvert\lambda\rvert \leq 1$	e_m	$\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda) \leq 2\sqrt{m}\,\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)$	$\begin{bmatrix} 0 & -A_3 & 0 \\ A_3 & A_2 & 0 \\ 0 & 0 & A_0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & A_3 \\ 0 & -A_3 & -A_2 \\ A_3 & A_2 & A_1 \end{bmatrix}$
odd	$\lvert\lambda\rvert \geq 1$	e_1	$\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda) \leq 2\sqrt{m}\,\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)$	$\begin{bmatrix} A_2 & A_1 & A_0 \\ -A_1 & -A_0 & 0 \\ A_0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} A_3 & 0 & 0 \\ 0 & A_1 & A_0 \\ 0 & -A_0 & 0 \end{bmatrix}$
even	$\lvert\lambda\rvert \geq 1$	e_1	$\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda) \leq 2\sqrt{m}\,\kappa_{\mathrm{P}}^{\mathbb{S}_e}(\lambda)$	$\begin{bmatrix} A_2 & 0 \\ 0 & A_0 \end{bmatrix} + \lambda \begin{bmatrix} A_1 & A_0 \\ -A_0 & 0 \end{bmatrix}$

Table 6.4: Recipes for choosing the ansatz vector v for a T-even or T-odd linearization L_e or L_o of a T-even matrix polynomial of degree m. Note that e_1 and e_m denote the 1st and mth unit vector of length m, respectively.

2. *If m is odd and $|\lambda| \geq 1$:* $\delta_{\Sigma v} \leq \dfrac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq 2\delta_{\Sigma v}.$

3. *If m is even and $|\lambda| \leq 1$:* $\delta_{\Sigma v} \leq \dfrac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \sqrt{2}\sqrt{m+1}\delta_{\Sigma v}.$

4. *If m is even and $|\lambda| \geq 1$:* $\delta_{\Sigma v} \leq \dfrac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq 2\delta_{\Sigma v}.$

In particular, considering the T-odd and T-even linearizations L_o and L_e corresponding to the ansatz vectors $v = \Pi_e(\overline{\Lambda}_{m-1})/\|\Pi_e(\Lambda_{m-1})\|_2$ and $v = (1-\Pi_e)(\overline{\Lambda}_{m-1})/\|(1-\Pi_e)(\Lambda_{m-1})\|_2$, respectively, we have $\delta_{\Sigma v} \leq \sqrt{2}$.

Proof:

1. Let m is odd and $|\lambda| \leq 1$. then Lemma 6.2.10 together with (6.28) yields

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} = \frac{\sqrt{1 - |y^T x|^2 \frac{1}{1+|\lambda|^2} \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^2}}}{\sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Since $\frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^2} \leq 1$ and $\frac{\|\Pi_e(\Lambda_m)\|_2}{\|\Lambda_m\|_2} = \frac{1}{1+|\lambda|^2}$ we have

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \geq \frac{\sqrt{1 - |y^T x|^2 \frac{1}{1+|\lambda|^2}}}{\sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \geq \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Further we have

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \frac{1}{\sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \cdot \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Now $\sqrt{1 - |y^T x|^2 \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1 - \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\|(I-\Pi_e)(\Lambda_m)\|_2}{\|\Lambda_m\|_2} \geq \frac{1}{\sqrt{m+1}}$. Hence the desired result follows by Lemma 6.4.3.

2. For m even $\frac{\|\Pi_e(\Lambda_m)\|_2}{\|\Lambda_m\|_2} \geq \frac{1}{1+|\lambda|^2}$ holds. Hence we have

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \geq \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

TH-789_04612301

Now $\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1-\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\|(I-\Pi_e)(\Lambda_m)\|_2}{\|\Lambda_m\|_2}$. Therefore we have

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \sqrt{2}\frac{\|\Lambda_m\|_2}{\|(I-\Pi_e)(\Lambda_m)\|_2}\delta_{\Sigma v}.$$

Now $\frac{\|\Lambda_m\|_2^2}{\|(I-\Pi_e)(\Lambda_m)\|_2^2} = 1+\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|(I-\Pi_e)(\Lambda_m)\|_2^2} = 1+\frac{1}{|\lambda|^2}$. Hence the result follows by Lemma 6.4.3.

3. If m is even then Lemma 6.2.10 together with (6.31) yields

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} = \frac{\sqrt{1-|y^Tx|^2\frac{1}{1+|\lambda|^2}\frac{|\Lambda_{m-1}^H\Sigma\Lambda_{m-1}|^2}{\|\Lambda_m\|_2^4}}}{\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \geq \frac{\sqrt{1-|y^Tx|^2\frac{1}{1+|\lambda|^2}}}{\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

It is easy to verify $\frac{1}{1+|\lambda|^2} \leq \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}$, implying

$$1 \leq \frac{\sqrt{1-|y^Tx|\frac{1}{1+|\lambda|^2}}}{\sqrt{1-|y^Tx|\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \Rightarrow \frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \geq \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Further

$$\frac{\kappa_{\mathrm{L}_o}^{\mathbb{S}_o}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \frac{1}{\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Now $\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1-\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\|(I-\Pi_e)(\Lambda_m)\|_2}{\|\Lambda_m\|_2} \geq \frac{1}{\sqrt{m+1}}$.
Therefore the result follows from Lemma 6.4.3.

4. For odd m, by considering a T-even linearization we obtain

$$\frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} = \frac{\sqrt{1-|y^Tx|\frac{|\lambda|^2}{1+|\lambda|^2}\frac{|\Lambda_{m-1}^T\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{1-|y^Tx|\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \geq \frac{\sqrt{1-|y^Tx|\frac{|\lambda|^2}{1+|\lambda|^2}}}{\sqrt{1-|y^Tx|\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

As above, it is not hard to verify $\frac{|\lambda|^2}{1+|\lambda|^2} \leq \frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}$, implying

$$1 \leq \frac{\sqrt{1-|y^Tx|\frac{|\lambda|^2}{1+|\lambda|^2}}}{\sqrt{1-|y^Tx|\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}} \Rightarrow \frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \geq \frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Further

$$\frac{\kappa_{\mathrm{L}_e}^{\mathbb{S}_e}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_o}(\lambda)} \leq \frac{1}{\sqrt{1-|y^Tx|\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}$$

Now $\sqrt{1-|y^Tx|^2\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} \geq \sqrt{1-\frac{\|\Pi_e(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}} = \frac{\|(I-\Pi_e)(\Lambda_m)\|_2}{\|\Lambda_m\|_2} \geq \frac{|\lambda|}{\sqrt{1+|\lambda|^2}}$.
Hence the result follows from Lemma 6.4.3 and (6.33) .

Finally, by (6.33) we have $\delta_{\Sigma v} \leq \sqrt{2}$. ■

TH-789_04612301

We mention that Table 6.4 has a virtually identical analogue in the case of a T-odd matrix polynomial.

Assuming that $\|\cdot\|_M \equiv \|\cdot\|_2$ the results discussed above present growth of structured condition number of a simple eigenvalue of P when P is linearized by a structured pencil in $\mathbb{L}_1(\mathrm{P})$. However, comparing $\kappa^{\mathbb{S}}_{\mathrm{L}}(\lambda)$ with $\kappa_{\mathrm{P}}(\lambda)$ we have the following results for $\|\cdot\|_M \equiv \|\cdot\|_F$ and $\|\cdot\|_M \equiv \|\cdot\|_2$.

Theorem 6.4.7. *Let $\mathbb{S}_e$ and $\mathbb{S}_o$ denote the sets of T-even and T-odd polynomials, respectively. Let λ be a finite simple eigenvalue of T-even polynomial* P. *Consider the T-even (resp. T-odd) linearizations L_e (resp. L_o) from $\mathbb{L}_1(\mathrm{P})$ corresponding to the ansatz vector $\Sigma v = v$ (resp. $\Sigma v = -v$). When $|\lambda| \leq 1$, we have the following:*

1. *For* $\|\cdot\|_M \equiv \|\cdot\|_F$: $\dfrac{\delta_{\Sigma v}}{\sqrt{2}} \leq \dfrac{\kappa^{\mathbb{S}_e}_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \delta_{\Sigma v}$ *and* $\dfrac{\sqrt{1-|x^Ty|^2}\,\|\Lambda_{m-1}\|_2}{\sqrt{2}\,\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}_o}_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \dfrac{\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v}$.

2. *For* $\|\cdot\|_M \equiv \|\cdot\|_2$: $\dfrac{\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}_e}_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \dfrac{\sqrt{2}\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v}$ *and* $\dfrac{\sqrt{1-|x^Ty|^2}\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}_o}_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \dfrac{\sqrt{2}\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v}$.

When $|\lambda| > 1$, we have the following.

1. *For* $\|\cdot\|_M \equiv \|\cdot\|_F$: $\dfrac{\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}_e}_{\mathrm{L}_e}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \delta_{\Sigma v}$ *and* $\dfrac{\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}_o}_{\mathrm{L}_o}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_{\Sigma v}$.

2. *For* $\|\cdot\|_M \equiv \|\cdot\|_2$ *and* $\mathrm{L} \in \{\mathrm{L}_e, \mathrm{L}_o\}$: $\dfrac{\sqrt{2-|x^Ty|^2}\|\Lambda_{m-1}\|_2}{\|\Lambda_m\|_2}\delta_{\Sigma v} \leq \dfrac{\kappa^{\mathbb{S}}_{\mathrm{L}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_{\Sigma v}, \mathbb{S} \in \{\mathbb{S}_e, \mathbb{S}_o\}$.

Proof: Note that both L_e and L_o preserve the eigensymmetry of P. First, consider the T-even linearization $\mathrm{L}_e \in \mathbb{L}_1(\mathrm{P})$. For $\|\cdot\|_M \equiv \|\cdot\|_F$, by Lemma 6.2.6, we have

$$\kappa^{\mathbb{S}}_{\mathrm{L}_e}(\lambda) = \frac{\sqrt{1+|\lambda|^2+(1-|\lambda|^2)|x^Ty|^2\frac{|\Lambda^H_{m-1}\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{2}}\frac{\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|\,\|\Lambda_m\|_2}\kappa_{\mathrm{P}}(\lambda).$$

Using (6.24) and noting that $\frac{|\Lambda^H_{m-1}\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4} \leq 1$ and $\|(1,\,\lambda)\|_2 \leq \sqrt{(1+|x^Ty|^2)+|\lambda|^2(1-|x^Ty|^2)} \leq \sqrt{2}$ when $|\lambda| \leq 1$, and

$$\sqrt{2} \leq \sqrt{(1+|x^Ty|^2\frac{|\Lambda^H_{m-1}\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4})+|\lambda|^2(1-|x^Ty|^2\frac{|\Lambda^H_{m-1}\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4})} \leq \|(1,\,\lambda)\|_2$$

when $|\lambda| > 1$, the desired results follow.

For $\|\cdot\|_M \equiv \|\cdot\|_2$, by Lemma 6.2.6 we have

$$\kappa^{\mathbb{S}}_{\mathrm{L}_e}(\lambda) = \frac{\sqrt{1+|\lambda|^2(1-|x^Ty|^2\frac{|\Lambda^H_{m-1}\Sigma\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4})}\,\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda,\Sigma v)|\,\|\Lambda_m\|_2}\kappa_{\mathrm{P}}(\lambda).$$

TH-789_04612301

Again using (6.24) and noting that

$$1 \leq \sqrt{1 + |\lambda|^2 (1 - |x^T y|^2 \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4})} \leq \sqrt{2}$$

when $|\lambda| \leq 1$, and

$$\sqrt{2 - |x^T y|^2} \leq \sqrt{1 + |\lambda|^2 (1 - |x^T y|^2 \frac{|\Lambda_{m-1}^H \Sigma \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4})} \leq \|(1, \lambda)\|_2$$

when $|\lambda| > 1$, the desired results follow. The proof is similar for T-odd linearization of P and follows from Lemma 6.2.10 ■

We mention that the results in Theorem 6.4.7 hold when P is T-odd matrix polynomial.

6.4.3 T-palindromic matrix polynomials

We now consider T-palindromic and T anti-palindromic matrix polynomials. T-palindromic and T-anti-palindromic polynomials put a restriction on the choice of the ansatz vector. Let $R \in \mathbb{R}^{m \times m}$ be given by

$$R = \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix} \tag{6.34}$$

Table 6.5 gives conditions for T-palindromic/T-anti-palindromic linearizations and can be found in [68, Tables 3.1]. The conditions in second and third columns of Table 6.5 are equivalent.

structure of P	(i) structure of $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$	(ii) $(R \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ and ansatz vector
$*$-palindromic	$*$-palindromic	$Rv = (v^*)^T$
	$*$-anti-palindromic	$Rv = -(v^*)^T$
$*$-anti-palindromic	$*$-palindromic	$Rv = -(v^*)^T$
	$*$-anti-palindromic	$Rv = (v^*)^T$

Table 6.5: Conditions the ansatz vector v needs to satisfy to yield a structured linearization such that $(R \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$.

Note that $\mathrm{P} \in \mathbb{S}$ has a structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ such that $(R \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$. Thus if $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ are the right and left eigenvectors of P corresponding to the eigenvalue λ then $\widetilde{x} = \Lambda_{m-1} \otimes x$ and $\widetilde{y} = R\overline{\Lambda_{m-1}} \otimes y$ are right and left eigenvectors corresponding to the eigenvalue λ of the structured linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$, respectively. Thus we have

$$\widetilde{y}^T \widetilde{x} = \frac{|\Lambda_{m-1}^H R \Lambda_{m-1}|}{\|\Lambda_{m-1}\|_2^2} y^T x \leq y^T x. \tag{6.35}$$

Proceeding in a similar way as that of Lemma 6.4.3 we have the following.

Lemma 6.4.8. *Let λ be a finite, simple eigenvalue of a matrix polynomial P with normalized right and left eigenvectors x and y. Then the eigenvalue condition number $\kappa_{\mathrm{L}}(\lambda)$ for a*

TH-789_04612301

linearization $\mathrm{L} \in \mathbb{L}_1(\mathrm{P})$ *with ansatz vector* v *such that* $(R \otimes I)\mathrm{L} \in \mathbb{DL}(\mathrm{P})$ *satisfies*

$$\kappa_{\mathrm{L}}(\lambda) = \frac{\sqrt{1+|\lambda|^2}}{|\mathsf{p}(\lambda; Rv)|} \cdot \frac{\|\Lambda_{m-1}\|_2^2}{|y^H P'(\lambda)x|} = \frac{\|(1,\,\lambda)\|_2 \|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda; Rv)|\, \|\Lambda_m\|_2} \kappa_{\mathrm{P}}(\lambda),$$

provided that the perturbations $\triangle \mathrm{L} = \triangle X + \lambda \triangle Y$ *are measured in the norm* $\|\triangle \mathrm{L}\| = \sqrt{\|\triangle X\|_M^2 + \|\triangle Y\|_M^2}$ *for a unitarily invariant norm* $\|\cdot\|_M$. *Thus, we have*

$$\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; Rv)|} \leq \frac{\kappa_{\mathrm{L}}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; Rv)|}.$$

Proof: The proof follows from Lemma 6.4.3, (6.35) and (6.24). ■

Note that $\delta_{Rv} = \dfrac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; Rv)|}$ serves as a growth factor for the unstructured condition number for a structured linearization $\mathrm{L}(\lambda) \in \mathbb{L}_1(\mathrm{P})$. Obviously $\delta_{Rv} \geq 1$. Note that $\mathsf{p}(\lambda; Rv) \neq 0$ and hence $\delta_{Rv} < \infty$ when L is structured linearization of P corresponding to v. For a finite simple eigenvalue of a T-palindromic matrix polynomial, we have the following.

Theorem 6.4.9. *Let* $\mathbb{S}_p$ *and* $\mathbb{S}_a$*, respectively, denote the sets of* T*-palindromic and* T*-anti-palindromic polynomials. Let* λ *be a finite simple eigenvalue of a* T*-palindromic matrix polynomial* P. *Let* L_p *(resp.* L_a*) be* T*-palindromic (resp.* T*-anti-palindromic) linearization from* $\mathbb{L}_1(\mathrm{P})$ *of* P *corresponding to the ansatz vectors* $v = Rv$ *(resp.* $v = -Rv$*). Then for* $\|\cdot\|_M \equiv \|\cdot\|_F$*, we have the following.*

1. *If* $\mathsf{re}(\lambda) > 0$: $\dfrac{\delta_{Rv}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_{Rv}$ *and* $\dfrac{|1-\lambda|\,\|\Lambda_{m-1}\|_2}{\sqrt{2}\|\Lambda_m\|_2}\delta_{Rv} \leq \dfrac{\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \delta_{Rv}$.
2. *If* $\mathsf{re} \leq 0$: $\dfrac{|1+\lambda|\,\|\Lambda_{m-1}\|_2}{\sqrt{2}\,\|\Lambda_m\|_2}\delta_{Rv} \leq \dfrac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \delta_{Rv}$ *and* $\dfrac{\delta_{Rv}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_{Rv}$.

Proof: First, consider L_p. By Lemma 6.2.11 we have

$$\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) = \frac{\|\Lambda_{m-1}\|_2^2 \sqrt{1+|\lambda|^2 + 2\mathsf{re}(\lambda)|y^T x|^2 \frac{|\Lambda_{m-1}^H R \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{2}\,|\mathsf{p}(\lambda; Rv)|\, \|\Lambda_m\|_2} \kappa_{\mathrm{P}}(\lambda).$$

If $\mathsf{re}(\lambda) > 0$ then

$$\|(1,\,\lambda)\|_2 \leq \sqrt{1+|\lambda|^2 + 2\mathsf{re}(\lambda)|y^T x|^2 \frac{|\Lambda_{m-1}^H R \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}} \leq \sqrt{2}\,\|(1,\,\lambda)\|_2.$$

Hence the result for L_p follows from Lemma 6.4.8 and (6.24). Similarly we obtain the result for $\mathsf{re}(\lambda) \leq 0$.

Next consider L_a. Then by Lemma 6.2.13

$$\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda) = \frac{\|\Lambda_{m-1}\|_2^2 \sqrt{1+|\lambda|^2 - 2\mathsf{re}(\lambda)|y^T x|^2 \frac{|\Lambda_{m-1}^H R \Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{2}\,|\mathsf{p}(\lambda; Rv)|\, \|\Lambda_m\|_2} \kappa_{\mathrm{P}}(\lambda).$$

TH-789_04612301

If $\mathsf{re}(\lambda) \leq 0$ then

$$\|(1,\,\lambda)\|_2 \leq \sqrt{1+|\lambda|^2 - 2\mathsf{re}(\lambda)|y^Tx|^2 \frac{\|\Lambda_{m-1}^H R\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}} \leq \sqrt{2}\,\|(1,\,\lambda)\|_2.$$

Hence the result for L_a follows from Lemma 6.4.8 and (6.24). Similarly we obtain the result for $\mathsf{re}(\lambda) > 0$. ■

In view of Table 6.5, the ansatz vector v for a T-palindromic linearization of a T-palindromic polynomial should satisfy $Rv = v$ with the flip permutation R defined in (6.34). By Proposition 6.4.4,

$$\max_{\substack{v=Rv\\ \|v\|_2\leq 1}} |\mathsf{p}(\lambda; Rv)| = \|\Pi_+(\Lambda_{m-1})\|_2,$$

where the maximum is attained by v_+ defined via

$$v_\pm = \frac{\left[\frac{\lambda^{m-1}\pm 1}{2}, \ldots, \frac{\lambda^{m/2+1}\pm\lambda^{m/2}}{2}, \frac{\lambda^{m/2+1}\pm\lambda^{m/2}}{2}, \ldots, \frac{\lambda^{m-1}\pm 1}{2}\right]^T}{\|\Pi_\pm(\Lambda_{m-1})\|_2} \tag{6.36}$$

if m is even and as

$$v_\pm = \frac{\left[\frac{\lambda^{m-1}\pm 1}{2}, \ldots, \frac{\lambda^{(m-1)/2}\pm\lambda^{(m-1)/2}}{2}, \ldots, \frac{\lambda^{m-1}\pm 1}{2}\right]^T}{\|\Pi_\pm(\Lambda_{m-1})\|_2} \tag{6.37}$$

if m is odd. Similarly,

$$\max_{\substack{v=-Rv\\ \|v\|_2\leq 1}} |\mathsf{p}(\lambda; Rv)| = \|\Pi_-(\Lambda_{m-1})\|_2,$$

with the maximum attained by v_-.

Theorem 6.4.10. *Let $\mathbb{S}_p$ and $\mathbb{S}_a$, respectively, denote the sets of T-palindromic and T-anti-palindromic polynomials. Let λ be a finite or infinite, simple eigenvalue of a* T-palindromic *matrix polynomial* P. *Let L_p (resp. L_a) be T-palindromic (resp. T-anti-palindromic) linearization from $\mathbb{L}_1(\mathrm{P})$ of* P *corresponding to the ansatz vectors $v = Rv$ (resp. $v = -Rv$). Then the following statements hold for $\|\cdot\|_M \equiv \|\cdot\|_F$.*

1. *If m is odd:* $\dfrac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} \leq \sqrt{2(m+1)}\,\delta_{Rv}$.

2. *If m is even and* $\mathsf{re}(\lambda) \geq 0$*:* $\dfrac{\delta_{Rv}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} \leq \sqrt{2(m+1)}\,\delta_{Rv}$.

3. *If m is even and* $\mathsf{re}(\lambda) \leq 0$*:* $\dfrac{\delta_{Rv}}{\sqrt{2}} \leq \dfrac{\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} \leq \sqrt{2(m+1)}\,\delta_{Rv}$.

Proof: Since $\kappa_{\mathrm{L}}^{\mathbb{S}}(\lambda)/\kappa_{\mathrm{P}}(\lambda) \leq \kappa_{\mathrm{L}}^{S}(\lambda)/\kappa_{\mathrm{P}}^{\mathbb{S}}(\lambda)$ holds for any structure $\mathbb{S}$, the desired lower bounds follow from Theorem 6.4.9.

1. If m is odd and $\mathsf{re}(\lambda) \geq 0$, Lemma 6.2.11 implies – together with Lemma 6.6.1.(1)

TH-789_04612301

and (6.28) – the inequality

$$\begin{aligned}\frac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} &= \frac{\sqrt{1-|y^Tx|^2\frac{|1+\lambda|^2-|1-\lambda|^2}{2(1+|\lambda|^2)}\frac{|\Lambda_{m-1}^H R\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{1-|y^Tx|^2\frac{\|\Pi_+(\Lambda_m)\|_2^2-\|\Pi_-(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \\ &\le \frac{1}{\sqrt{1-\frac{\|\Pi_+(\Lambda_m)\|_2^2-\|\Pi_-(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \\ &= \frac{\|\Lambda_m\|_2}{\sqrt{2}\|\Pi_+(\Lambda_m)\|_2}\cdot\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \le \sqrt{m+1}\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.\end{aligned}$$

Hence the desired result follows from Lemma 6.3.2.

2. The proof of the second part follows along the lines of the first part. For even m, Lemma 6.6.1(3) implies

$$\frac{\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} \le \frac{\|\Lambda_m\|_2}{\sqrt{2}\|\Pi_+(\Lambda_m)\|_2}\cdot\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \le \sqrt{m+1}\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

Hence the result follows from Lemma 6.3.2.

3. The proof of the third part also follows along the lines of the first part. Lemmas 6.2.11, 6.2.13 and 6.6.1(3) reveal – for even m and a T-anti-palindromic linearization – the inequality

$$\frac{\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)} = \frac{\sqrt{1-|y^Tx|^2\frac{|1-\lambda|^2-|1+\lambda|^2}{2(1+|\lambda|^2)}\frac{|\Lambda_{m-1}^H R\Lambda_{m-1}|^2}{\|\Lambda_{m-1}\|_2^4}}}{\sqrt{1-|y^Tx|\frac{\|\Pi_+(\Lambda_m)\|_2^2-\|\Pi_-(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2}}}\cdot\frac{\kappa_{\mathrm{L}_a}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \le \sqrt{m+1}\frac{\kappa_{\mathrm{L}_a}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)}.$$

The desired result follows from Lemma 6.3.2. ■

Corollary 6.4.11. *Let $\mathbb{S}_p$ and $\mathbb{S}_a$ denote the sets of T-palindromic and T-anti-palindromic polynomials, respectively. Let λ be a finite or infinite, simple eigenvalue of a T-*palindromic *matrix polynomial* P. *Consider the T-palindromic (resp. T-anti-palindromic) linearizations* $\mathrm{L}_p, \mathrm{L}_a \in \mathbb{L}_1(\mathrm{P})$ *belonging to the ansatz vectors v_+ and v_- defined in (6.36)–(6.37), respectively. Then the following statements hold for $\|\cdot\|_M \equiv \|\cdot\|_F$.*

1. *If m is odd: $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \le 2(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$.*
2. *If m is even and $\mathsf{re}(\lambda) \ge 0$: $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \le 2(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$.*
3. *If m is even and $\mathsf{re}(\lambda) \le 0$: $\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda) \le 2(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$.*

Proof: The desired results follow from Theorem 6.4.10 and Lemma 6.6.1. ■

Theorem 6.4.10 and Corollary 6.4.11 admit a simple interpretation. If either m is odd or m is even and λ has nonnegative real part, it is OK to use a T-palindromic linearization; there will be no significant (a small constant multiple of δ_{Rv}) increase of the structured condition number. In the exceptional case, when m is even and λ has negative real part, a T-anti-palindromic linearization should be preferred. This is especially true for $\lambda = -1$, in which case there is no T-palindromic linearization.

TH-789_04612301

The upper bounds in Corollary 6.4.10 are probably too pessimistic; at least they do not fully reflect the optimality of the choice of v_+ and v_-. In comparison, the heuristic choices listed in Table 6.6 yield almost the same bounds! These bounds are proven in the following lemma. To provide recipes for even m larger than 2, one would need to discriminate further between $|\lambda|$ close to 1 and $|\lambda|$ far away from 1, similar as for odd m.

m	λ of interest	v	Bound on struct. cond. of linearization	Example
odd	$\lvert\lambda\rvert \geq \alpha_m$ $\lvert\lambda\rvert \leq \alpha_m^{-1}$	$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$	$\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq 2\sqrt{2}(m+1)\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$	$\begin{bmatrix} A_0 & 0 & A_0 \\ A_1 - A_0^T & A_0 - A_1^T & 0 \\ A_1^T & A_1 - A_0^T & A_0 \end{bmatrix} + \lambda \begin{bmatrix} A_0^T & A_1^T - A_0 & A_1 \\ 0 & A_0^T - A_1 & A_1^T - A_0 \\ A_0^T & 0 & A_0^T \end{bmatrix}$
odd	$\lvert\lambda\rvert \leq \alpha_m$ $\lvert\lambda\rvert \geq \alpha_m^{-1}$	$e_{\frac{m-1}{2}}$	$\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq 2(m+1)\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$	$\begin{bmatrix} 0 & A_0 & 0 \\ 0 & A_1 & A_0 \\ -A_0^T & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & -A_0 \\ A_0^T & A_1^T & 0 \\ 0 & A_0^T & 0 \end{bmatrix}$
$m = 2$	$\mathsf{re}(\lambda) \geq 0$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq 2\sqrt{3}\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$	$\begin{bmatrix} A_0 & A_0 \\ A_1 - A_0^T & A_0 \end{bmatrix} + \lambda \begin{bmatrix} A_0^T & A_1^T - A_0 \\ A_0^T & A_0^T \end{bmatrix}$
$m = 2$	$\mathsf{re}(\lambda) \leq 0$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda) \leq 2\sqrt{3}\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$	$\begin{bmatrix} -A_0 & A_0 \\ -A_1 - A_0^T & -A_0 \end{bmatrix} + \lambda \begin{bmatrix} A_0^T & A_1^T + A_0 \\ -A_0^T & A_0^T \end{bmatrix}$

Table 6.6: Recipes for choosing the ansatz vector v for a T-palindromic or T-anti-palindromic linearization L_p or L_a of a T-palindromic matrix polynomial of degree m. Note that $\alpha_m = 2^{1/(m-1)}$.

Lemma 6.4.12. *The upper bounds on $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda)$ and $\kappa_{\mathrm{L}_a}^{\mathbb{S}_a}(\lambda)$ listed in Table 6.6 are valid.*

Proof: It suffices to derive an upper bound on $\frac{\|\Lambda_{m-1}\|_2}{\mathsf{p}(\lambda;v)}$. Multiplying such a bound by $\sqrt{2(m+1)}$ then gives the coefficient in the upper bound on the structured condition number of the linearization, see the proof of Theorem 6.4.10.

1. For odd m and $|\lambda| \geq \alpha_m$ or $|\lambda| \leq 1/\alpha_m$, the bound $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq \sqrt{2}(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$ follows from
$$\frac{\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|^2} \leq \frac{1 + \alpha_m^2 + \cdots + \alpha_m^{2m-2}}{|1 - \alpha_m^{m-1}|^2} = 1 + \alpha_m^2 + \cdots + \alpha_m^{2m-2} \leq 4m.$$

2. For odd m and $1/\alpha_m \leq |\lambda| \leq \alpha_m$, the bound $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq 2(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$ follows from
$$\frac{\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|^2} \leq \frac{1 + \alpha_m^2 + \cdots + \alpha_m^{2m-2}}{\alpha_m^{m-1}} = \frac{1}{2}(1 + \alpha_m^2 + \cdots + \alpha_m^{2m-2}) \leq 2m.$$

3. For $m = 2$ and $\mathsf{re}(\lambda) \geq 0$, the bound $\kappa_{\mathrm{L}_p}^{\mathbb{S}_p}(\lambda) \leq 2(m+1)\,\kappa_{\mathrm{P}}^{\mathbb{S}_p}(\lambda)$ follows for $|\lambda| \leq 1$ from
$$\frac{\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|^2} = \frac{1 + |\lambda|^2}{|1 + \lambda|^2} \leq 2$$
and for $|\lambda| \geq 1$ from
$$\frac{\|\Lambda_{m-1}\|_2^2}{|\mathsf{p}(\lambda;v)|^2} = \frac{|\lambda|^2}{|\lambda|^2}\,\frac{\frac{1}{|\lambda|^2} + 1}{|\frac{1}{\lambda} + 1|^2} \leq 2.$$

4. The proof for $m = 2$ and $\mathsf{re}(\lambda) \leq 0$ is analogous to Part 3.■

TH-789_04612301

For T-anti-palindromic polynomials, the results of Theorems 6.4.10 and 6.4.9, Corollary 6.4.11 and Table 6.6 hold, but with the roles of T-palindromic and T-anti-palindromic exchanged. For example, if either m is odd or m is even and $\mathsf{re}(\lambda) \geq 0$, there is always a good T-anti-palindromic linearization. Otherwise, if m is even and $\mathsf{re}(\lambda) \leq 0$, there is a good T-palindromic linearization.

6.4.4 H-Hermitian matrix polynomials and related structures

The linearization of a H-Hermitian polynomial is also H-Hermitian if the corresponding ansatz vector v is real, see Table 6.3. The optimal v, which maximizes $|\mathsf{p}(\lambda; v)|$, could be found by finding the maximal singular value and the corresponding left singular vector of the real $m \times 2$ matrix $[\mathsf{Re}(\Lambda_{m-1}), \mathsf{Im}(\Lambda_{m-1})]$. Instead of invoking the rather complicated expression for this optimal choice, the following lemma uses a heuristic choice of v.

Lemma 6.4.13. *Let $\mathbb{S}_h$ denote the set of H-Hermitian polynomials. Let λ be a finite or infinite, simple eigenvalue of a H-Hermitian matrix polynomial* P*. Then the following statements hold for $\|\cdot\|_M \equiv \|\cdot\|_F$.*

1. *If $|\lambda| \geq 1$ then the linearization* L *corresponding to the ansatz vector $v = [1, 0, \ldots, 0]$ is HHermitian and satisfies $\kappa_{\mathrm{L}}^{\mathbb{S}_h}(\lambda) \leq 2\sqrt{m}\kappa_{\mathrm{P}}^{\mathbb{S}_h}(\lambda)$.*

2. *If $|\lambda| \leq 1$ then the linearization* L *corresponding to the ansatz vector $v = [0, \ldots, 0, 1]$ is H-Hermitian and satisfies $\kappa_{\mathrm{L}}^{\mathbb{S}_h}(\lambda) \leq 2\sqrt{m}\kappa_{\mathrm{P}}^{\mathbb{S}_h}(\lambda)$.*

Proof: Assume $|\lambda| \geq 1$. Lemma 6.2.14 together with Lemma 6.3.2 and (6.24) imply

$$\frac{\kappa_{\mathrm{L}}^{\mathbb{S}_h}(\lambda)}{\kappa_{\mathrm{P}}^{\mathbb{S}_h}(\lambda)} \leq \sqrt{2}\,\frac{\kappa_{\mathrm{L}_p}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} = \sqrt{2}\,\frac{\|\Lambda_{m-1}\|_2}{|\mathsf{p}(\lambda; v)|} \leq 2\,\frac{\sqrt{m}|\lambda|^m}{|\lambda|^m} = 2\sqrt{m}.$$

The proof for $|\lambda| \leq 1$ proceeds analogously. ■

H-even and H-odd matrix polynomials are closely related to H-Hermitian matrix polynomials, see Remark 6.2.15. In particular, Lemma 6.4.13 applies verbatim to H-even and H-odd polynomials. Note, however, that in the case of even m the ansatz vector $v = [1, 0, \ldots, 0]$ yields an H-odd linearization for an H-even polynomial, and vice versa. Similarly, the recipes of Table 6.6 can be extended to H-palindromic polynomials.

Finally, comparing $\kappa_{\mathrm{L}}^{\mathbb{S}_h}(\lambda)$ with $\kappa_{\mathrm{P}}(\lambda)$, by Lemma 6.3.2 and (6.24) we have

$$\frac{\delta_v}{\sqrt{2}} \leq \frac{\kappa_{\mathrm{L}}^{\mathbb{S}_h}(\lambda)}{\kappa_{\mathrm{P}}(\lambda)} \leq \sqrt{2}\delta_v.$$

The same bound holds when P is H-even or H-odd and L is either H-even or H-odd linearization.

6.5 Summary and conclusions

We have derived relatively simple expressions for the structured eigenvalue condition numbers of certain structured matrix polynomials. These expressions have been used to analyze the possible increase of the condition numbers when the polynomial is replaced by a structured

TH-789_04612301

linearization. At least in the case when all coefficients of the polynomial are perturbed to the same extent, the result is very positive: There is always a structured linearization, which depends on the eigenvalue of interest, such that the condition numbers increase at most by a factor linearly depending on m. We have also provided recipes for structured linearizations, which do not depend on the exact value of the eigenvalue, for which the increase of the condition number is still negligible. Hence, the accuracy of a strongly backward stable eigensolver applied to the structured linearization will fully enjoy the benefits of structure on the sensitivity of an eigenvalue for the original matrix polynomial. The techniques and proofs of this chapter represent yet another testimonial for the versatility of the linearization spaces introduced by Mackey, Mackey, Mehl, and Mehrmann in [68, 67].

6.6 Appendix

The following lemma summarizes some auxiliary results needed in the proofs of Section 6.4.3.

Lemma 6.6.1. *Let $\lambda \in \mathbb{C}$ and let $\Lambda_\pm$ be defined as in (6.23. Then the following statements hold.*

1. *If m is odd and $\mathsf{re}(\lambda) \geq 0$ then $\frac{\|\Pi_+(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \geq \frac{1}{2(m+1)}$.*
2. *If m is odd and $\mathsf{re}(\lambda) \leq 0$ then $\frac{\|\Pi_+(\Lambda_m)\|_2^2 - \|\Pi_-(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \geq \frac{|1+\lambda|^2 - |1-\lambda|^2}{2(1+|\lambda|^2)}$.*
3. *If m is even then $\frac{\|\Pi_+(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \geq \frac{1}{2(m+1)}$.*
4. *If m is odd and $\mathsf{re}(\lambda) \leq 0$ then $\frac{\|\Pi_-(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} \geq \frac{1}{2(m+1)}$.*

Proof:

1. For $|\lambda| \geq 1$ the statement follows from $\|\Lambda_m\|_2^2 \leq (m+1)|\lambda|^{2m}$ and

$$\begin{aligned} 2\|\Pi_+(\Lambda_m)\|_2^2 &\geq |\lambda^m + 1|^2 + |\lambda^{(m+1)/2} + \lambda^{(m-1)/2}|^2 \\ &= |\lambda|^{2m} + 2\mathsf{re}(\lambda^m) + 1 + |\lambda|^{m-1}(|\lambda|^2 + \mathsf{re}(\lambda) + 1) \\ &\geq |\lambda|^{2m} - 2|\lambda|^m + 1 + |\lambda|^{m-1}(|\lambda|^2 + 1) \\ &= |\lambda|^{2m} + 1 + |\lambda|^{m-1}(|\lambda| - 1)^2 \geq |\lambda|^{2m}. \end{aligned}$$

   If $|\lambda| \leq 1$, we can apply an analogous argument with λ replaced by $1/\lambda$ to

$$\frac{\|\Pi_+(\Lambda_m)\|_2^2}{\|\Lambda_m\|_2^2} = \frac{|\lambda|^{2m}}{|\lambda|^{2m}} \cdot \left(\sum_{k=0}^{(m-1)/2} \frac{1}{2} \left| \frac{1}{\lambda^k} + \frac{1}{\lambda^{m-k}} \right|^2 \right) \Big/ \left(\sum_{k=0}^{m} \frac{1}{|\lambda|^{2k}} \right).$$

2. Using $(1+|\lambda|^2)(1+|\lambda|^4+\cdots+|\lambda|^{(2m-2)}) = \|\Lambda_m\|_2^2$, we prove the equivalent statement

$$\|\Pi_+(\Lambda_m)\|_2^2 \geq (|1+\lambda|^2 - |1-\lambda|^2)(1+|\lambda|^4+\cdots+|\lambda|^{(2m-2)}).$$

   Assume $|\lambda| \leq 1$. Then the statement follows if we can show

$$\|\Pi_+(\Lambda_m)\|_2^2 - \|\Pi_-(\Lambda_m)\|_2^2 \geq \frac{1}{4}(|1+\lambda|^2 - |1-\lambda|^2)(m+1). \tag{6.38}$$

TH-789_04612301

Inserting $\lambda = |\lambda|(\cos(\phi) + \mathrm{i}\sin(\phi))$, we expand

$$\begin{aligned}
\|\Pi_+(\Lambda_m)\|_2^2 - \|\Pi_-(\Lambda_m)\|_2^2 &= \frac{1}{2}\sum_{k=0}^{(m-1)/2} |\lambda^{m-k} + \lambda^k|^2 - |\lambda^{m-k} - \lambda^k|^2 \\
&= 2|\lambda|^m \sum_{k=0}^{(m-1)/2} (\cos((m-k)\phi)\cos(k\phi) + \sin((m-k)\phi)\sin(k\phi)) \\
&= 2|\lambda|^m \sum_{k=0}^{(m-1)/2} \cos((m-2k)\phi) = 2|\lambda|^m \sum_{k=0}^{(m-1)/2} \cos(\phi + 2k\phi) \\
&= 2|\lambda|^m \frac{\sin(\frac{m+1}{2}\phi)\,\cos(\frac{m+1}{2}\phi)}{\sin\phi} = 2|\lambda|^m \frac{\sin((m+1)\phi)}{\sin\phi}.
\end{aligned}$$

On the other hand,

$$|1+\lambda|^2 - |1-\lambda|^2 = 4|\lambda|\frac{\sin(2\phi)}{\sin\phi}.$$

Thus (6.38) is equivalent to

$$|\lambda|^{m-1}\frac{\sin((m+1)\phi)}{\sin\phi} \geq \frac{m+1}{2}\frac{\sin(2\phi)}{\sin\phi}$$

Dividing by $\cos(\phi) \leq 0$ on both sides, this is in turn equivalent to

$$|\lambda|^{m-1}\frac{\sin((m+1)\phi)}{\sin(2\phi)} \leq \frac{m+1}{2}.$$

Finally, using $|\lambda| \leq 1$, the last inequality follows from the basic trigonometric inequality $\frac{\sin((m+1)\phi)}{\sin(2\phi)} \leq \frac{m+1}{2}$. For $|\lambda| \geq 1$, we can use the same trick as in the proof of part 1 and replace λ by $1/\lambda$.

3. As in part 1, we can assume w.l.o.g. $|\lambda| \geq 1$. Then $\|\Lambda_m\|_2^2 \leq (m+1)|\lambda|^{2m}$ and

$$\begin{aligned}
2\|\Pi_+(\Lambda_m)\|_2^2 &\geq |\lambda^m + 1|^2 + 2|\lambda|^m \\
&= |\lambda|^{2m} + 2\mathsf{re}(\lambda^m) + 1 + 2|\lambda|^m \geq |\lambda|^{2m},
\end{aligned}$$

concluding the proof.

4. This part follows from Part 1, by simply replacing $\lambda \to -\lambda$, which implies $\|\Pi(\Lambda_+)\|_2 \leftrightarrow \|\Pi(\Lambda_-)\|_2$.

■

TH-789_04612301

Conclusion

We have undertaken a detailed structured backward perturbation and sensitivity analysis of structured polynomial eigenvalue problem including complex symmetric, skew-symmetric, Hermitian, even, odd, palindromic and anti-palindromic problem.

First, we have provided a complete solution of structured mapping problem for matrices. We have also provided a complete solution of structured inverse least squared problem for matrices. We have shown that these results play an important role in determining structured backward errors of approximate invariant subspaces of structured matrices. With the help of these results, we have analyzed structured pseudospectra of structured matrices.

Second, we have analyzed structured backward perturbation of structured matrix polynomials. We have determined structured backward error of an approximate eigenelement of a structured matrix polynomial and have determined a minimal structured perturbation such that the approximate eigenelement is the exact eigenelement of the perturbed problem. We have also analyzed structured pseudospectra of structure matrix polynomials. Further, we have analyzed structured condition numbers of simple eigenvalues of structured matrix polynomials and have derived explicit expressions for the condition numbers. Structured condition numbers measure the sensitivity of simple eigenvalues to small structured perturbations and hence play an important role in the accuracy assessment of approximate eigenelements of structured polynomial eigenvalue problem.

Finally, most numerical methods for solving polynomial eigenvalue problem proceed by linearizing the polynomial eigenvalue problem into an equivalent generalized eigenvalue problem of larger size and solve the resulting problem. Therefore, for computational purposes, it is of paramount importance to identify potential structured linearizations which are as well conditioned as possible. With the help of structured backward perturbation analysis and structured condition numbers of eigenvalues, we have provided a recipe for identifying "good" structured linearizations which guarantee almost as accurate solutions as that of the original polynomial eigenvalue problem.

TH-789_04612301

Bibliography

[1] S. S. Ahmad, *Pseudospectra of matrix pencils and their applications in perturbation analysis of eigenvalues and eigendecompositions,* PhD thesis, Department of Mathematics, Indian Institute of Technology Guwahati, India, 2007.

[2] S. S. Ahmad and R. Alam, *Pseudospectra, critical points and multiple eigenvalues of matrix polynomials,* Linear Algebra Appl. (2008), doi:10.1016/j.laa.2008.10.019.

[3] S. S. Ahmad, R. Alam and R. Byers, *Pesudospectra, critical points and multiple eigenvalues of matrix pencils,* submitted.

[4] A. L. Andrew, E. K.-W. Chu, and P. Lancaster, *Derivatives of eigenvalues and eigenvectors of matrix functions,* SIAM J. Matrix Anal. Appl., 14(1993) pp.903-926.

[5] B. Aupetit, *A Primer on Spectral Theory,* Springer-Verlag, 1991.

[6] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templets for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, 2000.

[7] G. P. Barker, *Monotone Norms,* Nuer. Math. 18(1972) pp.321-326.

[8] F. S. V. Bazán, *Eigenvalues of Matrix Polynomials, Sensitivity, Computation and Applications* (in Portuguese). IMPA, Rio de Janeiro, Brazil, 2003.

[9] 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(1998), pp.329-358.

[10] P. Benner, V. Mehrmann and H. Xu, *A note on the numerical solution of complex Hamiltonian and skew-Hamiltonian eigenvalue problems,* Elect. Trans. Numer. Anal., 8(1999), pp. 115-126.

[11] R. Bhatia, *Matrix Analysis,* Springer-Verlag, New York, 1997.

[12] R. Bhatia, M.-D. Choi and C. Davis, *Comparing a matrix to its off-diagonal part,* Oper. Theory Adv. Appl., 40(1989), pp. 151-164.

[13] R. Bhatia and F. Kittaneh, *Commutators, pinchings and spectral variation,* Oper. Matrices, 2(2008), pp. 143-151.

[14] M. Berhanu, *The polynomial eigenvalue problem,* PhD thesis, School of Mathematics, The University of Manchester, UK., 2005.

[15] T. Betke, N. Higham, V. Mehrmann, C. Schröder and F. Tisseur, NLEVP:*A collection of nonlinear eigenvalue problems,* MIMS eprint:2008.40, April, 2008.

[16] S. Bora, *Structured eigenvalue conditioning and backward error of a class of polynomial eigenvalue problems,* Preprint 2009.

TH-789_04612301

[17] L. Boulton, P. Lancaster and P. Psarrakos, *On pseudospectra of matrix polynomials and their boundaries*, Math. Comp., Electronically published, May 11, 2007.

[18] A. Bunse-Gerstner, R. Byers and V. Mehrmann, *A chart of numerical methods for structured eigenvalue problems,* SIAM J. Matrix Anal. Appl., 13(1992), pp. 419-453.

[19] R. Byers and D. Kressner, *On the condition of a complex eigenvalue under real perturbations,* BIT, 44(2004) pp.209-215.

[20] R. Byers and D. Kressner, *Structured condition numbers for invariant subspaces,* SIAM J. Matrix Anal. Appl., 28(2006) pp.326-347.

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

[22] E. Chu, T. Hwang, W. Lin and C. Wu, *Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms,* J. Comp. Appl. Math., 219(2008), pp. 237-252.

[23] C. Davis, *A factorization of an arbitrary $m \times n$ contractive operator-matrix,* Toeplitz centennial (I. Gohberg, Ed.), Basel: Birkhuser, pp. 217-232, 1982.

[24] C. Davis, W. M. Kahan and H.F. Weinberger, *Norm-preserving dialations and their applications to optimal error bounds,* SIAM J. Numer. Anal., 19(1982) 445-469.

[25] J. P. Dedieu, *Condition operators, condition numbers, and condition number theorem for generalized eigenvalue problem*, Linear Algebra Appl., 263 (1997), pp. 1–24.

[26] J.-P. Dedieu and F. Tisseur, *Perturbation theory for homogeneous polynomial eigenvalue problems,* Linear Algebra Appl., 358(2003), pp.71-94

[27] J. W. Demmel and B. Kågström, *Computing stable eigendecompositions of matrix pencils*, Linear Algebra Appl., 88–89 (1987), pp. 139–186.

[28] M. Embree and L. N. Trefethen, Pseudospectra Gateway. Available from: <http://web.comlab.ox.ac.uk/pseudospectra> .

[29] H.-Y. Fan, W.-W. Lin and P. Van Dooren, *Normwise scaling of second order polynomial matrices,* SIAM J. Matrix Anal. Appl., 26(2004), pp.252-256.

[30] R.W. Freund and F. Jarre, *An extension of the positive real lemma to descriptor systems,* Optimization methods and software, 19(2004), pp.69-87.

[31] F. R. Gantmacher, *The Theory of Matrices,* Vols. 1, 2. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959.

[32] I. Gohberb, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press, 1982.

[33] M. Golomb and H. F. Weinberger, *Optimal approximation and error bounds,* On numerical approximation (R.E. Langer, Ed.), Madison : The University of Wisconsin Press, 1959,pp. 117190.

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

[35] S. Graillat, *A note on structured pseudospectra,* J. Comp. Appl. Math., 191(2006) pp. 68-76.

TH-789_04612301

[36] N. J. Higham, *Accuracy and Stability of Numerical Algorithms,* SIAM, Philadelphia, PA, second edition, 2002.

[37] D. J. Higham and N. J. Higham, *Backward error and condition of structured linear systems,* SIAM J. Matrix Anal. Appl., 13 (1992), pp.162-175.

[38] D. J. Higham and N. J. Higham, *Structured backward error and condition of generalized eigenvalue problems,* SIAM J. Matrix Anal. Appl., 20(1998), pp.493-512.

[39] N. J. Higham, R. Li, and F. Tisseur, *Backard error of polynomial eigenproblems solved by linearization,* SIAM J. Matrix Anal. Appl., 29(2007), pp. 1218 - 1241.

[40] N. J. Higham, D. S. Mackey, N. Mackey and F. Tisseur, *Symmetric linearizations for matrix polynomials,* SIAM J. Matrix Anal. Appl., 29(2006), pp. 143 - 159.

[41] N. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials,* SIAM J. Matrix Anal. Appl., 28(2006), pp. 1005 - 1028.

[42] N. J. Higham, D. S. Mackey, F. Tisseur and S. D. Garvey, *Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problem.,* Internat. J. Numer. Methods Engrg., 37(2008), pp.344-360.

[43] N. J. Higham and F. Tisseur, *More on pseudospectra for polynomial eigenvalue problems and applications in control theory,* Linear Algebra Appl., 351-352 (2002), pp. 435–453.

[44] A. Hilliges, C. Mehl, and V. Mehrmann, *On the solution of palindromic eigenvalue problems,* In Proceedings of ECCOMAS, Jyväskylä, Finland, 2004.

[45] R.A. Horn and C. R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, reprinted paperback version ed., 1995.

[46] T.-M. Hwang, W.-W. Lin and V. Mehrmann, *Numerical solution of quadratic eigenvalue problems with structure-preserving methods,* SIAM J. Scientific Comp., 24(2003), pp.1283-1302.

[47] B. Kågström and P. Poromaa, *Computing eigenspaces with specified eigenvalues of a regular matrix pair (A,B) and condition estimation: Theory, algorithms and software,* Numer. Algorithms, 12(1996) pp.369-407.

[48] W. Kahan, B. N. Parlett and E. Jiang, *Residual bounds on approximate eigensystems of nonnormal matrices,* SIAM J. Numer. Anal. 19 (1982), pp.470-484.

[49] M. Karow, *μ-values and spectral value sets for linear perturbation classes defined by a scalar product,* Submitted.

[50] M. Karow, *Structured pseudospectra and the condition of a nonderogatory eigenvalue,* Submitted 2007.

[51] M. Karow, D. Kressner and F. Tisseur, *Structured eigenvalue condition numbers,* SIAM J. Matrix Anal. Appl., 28(2006), pp.1052-1068.

[52] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[53] D. Kressner, *Numerical Methods for General and Structured Eigenvalue Problems,* volume 46 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2005.

[54] D. Kressner, M. J. Pelaez and J. Moro, *Structured Hölder condition numbers for multiple eigenvalues,* Uminf report, Department of Computing Science, Umeȧ University, Sweden, October 2006.

TH-789_04612301

[55] D. Kressner, C. Schröder and D. S. Watkins, *Implicit QR algorithms for palindromic and even eigenvalue problems,* Preprint 432, DFG Research Center Matheon, Mathematics for key technologies in Berlin, TU Berlin, Germany, January 2008.

[56] A. Kurzhanski and I. Valyi, *Ellipsoidal calculus for estimation and control,* Systems and Control: Foundations and Applications. Birkhauser Boston Inc., Boston, MA, 1997.

[57] P. Lancaster, *Lambda m-Matrices and Vibrating Systems*, Dover Publication, 1966.

[58] P. Lancaster, A. S. Markus, and F. Zhou, *Perturbation theory for analytic matrix functions: the semisimple case,* SIAM J. Matrix Anal. Appl., 25(2003) pp. 606-626.

[59] P. Lancaster and P. Psarrakos, *On the pseudospectra of matrix polynomials,* SIAM J. Matrix Anal. Appl., 27 (2005), pp.115–129.

[60] H. Langer and B. Najman, *Remarks on the perturbation of analytic matrix functions. II,* Integral Equations Operator Theory, 12(1989) pp.392-407.

[61] H. Langer and B. Najman, *Leading coecients of the eigenvalues of perturbed analytic matrix functions,* Integral Equations Operator Theory, 16(1993) pp.600-604.

[62] Xin-guo Liu and Wei-guo Wang, *On a class of alternating coefficient matrices quadratic eigenvalue problem,* Applied Mathematics and Computation, 158(2004) pp.619-636.

[63] Xin-guo Liu and Ze-xi Wang, *A note on the backward errors for Hermite eigenvalue problems,* Applied Mathematics And Computation, 165(2005), pp.405-417.

[64] D. S. Mackey, *Structured linearizations for matrix polynomials,* PhD thesis, School of Mathematics, The University of Manchester, UK, 2006.

[65] D. S. Mackey, N. Mackey and D. Dunlavy, *Structure preserving algorithms for perplectic eigenproblems,* Elect. J. Linear Algebra, 13(2005), pp.10-39.

[66] D. S. Mackey, N. Mackey, C. Mehl and V. Mehrmann, *Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form,* To appear in Numer. Linear Algebra Appl.

[67] D. S. Mackey, N. Mackey, C. Mehl and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials,* SIAM J. Matrix Anal. Appl., 28(2006), pp.971-1004.

[68] D. S. Mackey, N. Mackey, C. Mehl and V. Mehrmann, *Palindromic polynomial eigenvalue problems: good vibrations from good linearizations,* SIAM J. Matrix Anal. Appl., 28(2006), pp.1029-1051.

[69] D. S. Mackey, N. Mackey and F. Tisseur, *Structured tools for structured matrices,* Elect. J. Linear Algebra, 10(2003): pp.106-145.

[70] D. S. Mackey, N. Mackey and F. Tisseur, *Structured mapping problems for matrices associated with scalar products, Part I: Lie and Jordan algebras,* SIAM J. Matrix Anal. Appl., 29(2008), pp.1389-1410.

[71] D. S. Mackey, N. Mackey and F. Tisseur, *Structured factorizations in scalar product spaces,* SIAM J. Matrix Anal. Appl., 27(2006), pp.821-850.

TH-789_04612301

[72] V. Mehrmann and H. Xu, *Structured Jordan canonical forms for structured matrices that are Hermitian, skew-Hermitian or unitary with respect to indefinite inner products,* Elect. J. Linear Algebra, 5(1999), pp.67-103.

[73] V. Mehrmann and H. Xu, *Perturbation of purely imaginary eigenvalues of Hamiltonian matrices under structured perturbations,* Elect. J. Linear Algebra, 17(2008) pp.234-257.

[74] V. Mehrmann and D. Watkins, *Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils,* SIAM J. Scientific Comp., 22(2001), pp.1905-1925.

[75] V. Mehrmann and D. Watkins, *Polynomial eigenvalue problems with Hamiltonian structure,* Elect. Trans. Numer. Anal., 13(2002), pp.106-113.

[76] J. Meinguet, *On the Davis-Kahan-Weinberger solution of the norm-preserving dilation problem,* Numer. Math., 49(1986), pp. 331-341.

[77] C. A. Micchelli and T. J. Rivlin, *A survey of optimal recovery,* Optimal estimation in approximation theory (C.A. Micchelli and T.J. Rivlin, Eds), New York : Plenum Press 1977, pp. 1-54

[78] S. Noschese and L. Pasquini, *Eigenvalue condition numbers: zero-structured versus traditional,* J. Comput. Appl. Math., 185(2006) pp.174-189.

[79] S. M. Rump, *Structured perturbations Part I: normwise distances,* SIAM J. Matrix Anal. Appl., 25(2003) pp.1-30.

[80] S. M. Rump, *Structured perturbations Part II: componentwise distances,* SIAM J. Matrix Anal. Appl., 25(2003), pp. 31-56.

[81] S. M. Rump, *Eigenvalues, pseudospectrum and structured perturbation,* Linear Algebra Appl., 413(2006), pp.567-593.

[82] C. Schröder, Palindromic and even eigenvalue problems - analysis and numerical methods, PhD thesis, Technical University Berlin, Germany, 2008.

[83] J. Simo, K. Wong, *Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum,* Internat. J. Numer. Methods Engrg., 31(1991), pp.19-52.

[84] Willi-Hans Steeb, *Matrix Calculas and Kronecker Product with Applications and* C^{++} *Programs,* World Scientific Publishing Co., NJ, USA, 1997.

[85] G. W. Stewart, *On the sensitivity of the eigenvalue problem* $Ax = \lambda Bx$, SIAM J. Numer. Anal. Appl., 9 (1972), pp. 669–686.

[86] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems,* SIAM Rev., 15(1973), pp. 727-764.

[87] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, Boston,1990.

[88] J.-G. Sun, *Stability and accuracy: perturbation analysis of algebraic eigenproblems* Technical Report UMINF 98-07, Department of Computing Science, University of Umeå, Umeå, Sweden, 1998.

[89] J. -G. Sun, *Least-squares solutions of a class of inverse eigenvalue problems for real symmetric matrices (in Chinese),* Math. Numer. Sin., 3(1988) , pp. 282-290.

TH-789_04612301

[90] J. -G. SUN, *Two kinds of inverse eigenvalue problems(in Chinese),* Math. Numer. Sin., 9(1987) , pp. 206-216.

[91] J. SUN, *Backward perturbation analysis of certain characteristic subspaces,* Numer. Math., 65 (1993), pp. 357-382.

[92] R. C. THOMPSON, *Pencils of complex and real symmetric and skew matrices,* Linear Algebra Appl., 147(1991) pp.323371.

[93] F. TISSEUR, *Backward error and condition of polynomial eigenvalue problems,* Linear Algebra Appl., 309(2000), pp.339-361.

[94] F. TISSEUR, *Stability of structured Hamiltonian eigensolvers,* SIAM J. Matrix Anal. Appl., 23(2001), pp.103-125.

[95] F. TISSEUR, *A chart of backward errors and condition numbers for singly and doubly structured eigenvalue problems,* SIAM J. Matrix Anal. Appl., 24(2003), pp.877-897.

[96] F. TISSEUR AND S. GRAILLAT, *Structured condition numbers and backward errors in scalar product spaces,* Elect. J. Linear Algebra, 15: 159-177, 2006.

[97] F. TISSEUR AND N. J. HIGHAM, *Structured pseudospectra for polynomial eigenvalue problems, with applications,* SIAM J. Matrix Anal. Appl. 23(2001), pp.187-208.

[98] F. TISSEUR, D. S. MACKEY AND N. MACKEY, *G-refectors: analogues of Householder transformations in scalar product spaces,* Linear Algebra Appl., 385(2004) pp.187-213.

[99] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem,* SIAM Rev. 43(2001), pp.235-286.

[100] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: the Behaviour of Nonnormal Matrices and Operators,* Princeton University Press, 2005.

[101] G. TRENKLER, *Matrices which take a given vector into a given vector revisited,* Amer. Math. Monthly, 111 (2004), pp.50-52.

[102] H. F. WEINBERGER, *On optimal numerical solution of partial differential equations* SIAM J. Numer. Anal.,9(1972), pp.182-198.

[103] J. H. WILKINSON, *Sensitivity of eigenvalues*, Util. Math., 25 (1984), pp. 5–76.

[104] J. H. WILKINSON, *Sensitivity of eigenvalues II*, Util. Math., 30 (1986), pp. 243–286.

[105] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, UK, 1996.

[106] D. X. XIE, L. ZHANG AND X. Y. HU, *Least-squares solutions of inverse problems for bisymmetric matrices(in Chinese),* Math. Numer. Sin., 22(2000), pp. 29-40.

[107] Z. ZHANG AND C. LIU, *Least-squares solutions of the equation $AX = B$ over anti-Hermitian generalized Hamiltonian matrices,* Numer. Maths., A Journal of Chinese Universities(English Series), 15(2006), pp. 60-66.

TH-789_04612301

List of Publications/Communicated

1. B. Adhikari and R. Alam, *Structured backward errors and pseudospectra of structured matrix pencils,* SIAM J. Matrix Anal. Appl., 31(2009), pp.331-359.

2. B. Adhikari, *Backward errors and linearizations for palindromic matrix polynomials,* preprint arXiv0812.4154 math.NA Submitted to Linear Algebra Appl..

3. B. Adhikari, R. Alam and D. Kressner, *Structured eigenvalue condition numbers and linearizations for matrix polynomials,* preprint available at http://www.math.ethz.ch/kressner/pubs.php Submitted to Linear Algebra Appl..

4. B. Adhikari and R. Alam, *Structured backward errors and pseudospectra of structured matrix polynomials,* to be submitted.

5. B. Adhikari and R. Alam, *Structured mapping problems for linearly structured matrices,* to be submitted.

TH-789_04612301