Structured matrices in indefinite inner product spaces: simple forms, invariant subspaces and rank-one perturbations.

D.B. JANSE VAN RENSBURG

10838368

Thesis submitted for the degree Philosophia Doctor of Natural Science at the Potchefstroom Campus of the North-West University.

Promoter:	**Prof. G.J. Groenewald**
Co-promoter:	**Prof. J.H. Fourie**
Assistant-promoter:	**Prof. A.C.M Ran**

April 2012

Keywords: H-expansive matrices, H-positive real matrices, invariant semi definite subspaces, indefinite inner product spaces, rank-one perturbations

Voorwoord

Met die bereiking van 'n mylpaal - soos die afhandeling van 'n tesis - besef mens altyd dat baie spesiale mense belangrike insette gelewer het. Nie net op akademiese gebied nie, maar in elke aspek van menswees. Die liefde vir wiskunde kom al van kindsbeen af, daar op die plaas in Ermelo, waar skape gereeld getel moes word en daar ook op een of ander manier 'n wiskunde sommetjie gedoen moes word. Tafels leer was van kardinale belang. Op hoërskool is my wiskundige vaardigheid bevorder deur iemand wat 'n liefde daarvoor het - dankie Lenie.

Die betowering van die noue wisselwerking tussen wiskunde en die praktyk en die liefde vir onderrig het gelei tot 'n loopbaan aan die NWU. Na 'n vyf jaar studie tydperk aan die NWU was daar eers die verdere persoonlike ontwikkeling met 'n studie aan die universiteit van Eindhoven, Nederland. Daar het ek nie net baie van wiskunde geleer nie, maar ook van myself. Behalwe vir die wiskunde het my liefde vir hardloop en tennis my ook in kontak gebring met die kultuur en omgewing van 'n vreemde land, mense en plekke.

Terug in my geboorteland het my paadjie eers gelei na my oud-skool vir een jaar, alvorens ek as junior lektor aan die NWU aangestel is. Hier het ek my nessie begin skrop en Potchefstroom my tuiste gemaak.

Na 'n paar jaar van voete vind, het 'n PhD studie op my gewag. Met soveel moontlike rigtings was dit moeilik om te kies. Maar deur noue samewerking van my studieleiers, Gilbert Groenewald en Jan Fourie en met Andre Ran van Amsterdam (NL), was dit gou duidelik dat hier 'n wen kombinasie is. Tydens 'n besoek van Andre Ran aan die NWU is daar op 'n PhD projek besluit. Gou het ek my leemte in baie aspekte van die wiskunde ontdek, maar deur goeie leiding van Gilbert, Jan en veral Andre was daar altyd vordering gewees. Dit was soms stoei en kopkrap, maar die oplossings was te vinde.

Deur my navorsing saam met Andre het ek in kontak gekom met groot name in die gebied van Lineêre Algebra, Operator Teorie en ander gebiede van Wiskunde. Met die bywoning van internasionale kongresse en werkswinkels, het ek maar net 'n klein blik gekry in hoe groot en wyd Wiskunde werklik is.

Die eise van navorsing en onderrig het nie my liefde vir sport gedemp nie, trouens, dit was altyd 'n welkome ontvlugting. Op die pad, fiets of in die swembad is baie keer nagedink oor probleme en die soeke na moontlike oplossings. Maar daar was ook tyd vir ontspanning saam met vriende, wat my soveel ondersteuning gegee het.

Bo dit alles, was dit altyd my Skepper en Verlosser wat my vasgehou en deurgedra het. Behalwe vir Hom, wil ek net dankie sê vir my familie se getroue ondersteuning - veral my ouers se gebede en harde werk wat dit moontlik gemaak het vir my om te studeer en my susters se liefde en bystand deur dik en dun.

Summary

The (definite) inner product between two vectors $x, y \in \mathbb{R}^n$ is defined by $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$. The length of a vector $x \in \mathbb{R}^n$ is then described by the inner product as $\|x\| = \langle x, x\rangle^{1/2} = \sqrt{x_1^2 + \ldots + x_n^2}$. In this thesis the definite inner product is replaced by an indefinite inner product. This has a substantial impact on the geometry of subspaces. The indefinite inner product between two vectors in $\mathbb{R}^n$ can uniquely be represented by $[x, y] = \langle Hx, y\rangle$, for a real invertible symmetric matrix $H = H^T$.

In the literature there exists an extensive theory for the classes of hermitian, unitary and normal matrices for the definite inner product space. Similar classes are also studied in the context of indefinite inner product space. In this thesis we focus on a study of H-positive real and H-expansive matrices. This study is closely related to a thorough study of H-dissipative and H-contractive matrices, which already exists in the literature.

In Chapter 2 we discuss a class of matrices that is closely related to H-dissipative matrices, namely, the class of H-positive real matrices. A matrix A is H-dissipative if $\frac{1}{2i}(HA - A^*H) \geq 0$. An $n \times n$ complex matrix A is H-positive real if $\frac{1}{2}(HA + A^*H) \geq 0$. It is easily seen that A is H-positive real if and only if iA is H-dissipative, and hence it follows that $-iA$ is H-positive real if and only if A is H-dissipative.

A simple form for the matrix H is obtained and thereafter the A-invariant maximal H-nonnegative and H-nonpositive subspaces are constructed. The main focus however, is to obtain the uniqueness and stability of these subspaces. The numerical range condition is used for this purpose. Both the real and complex cases are treated in this chapter.

In Chapter 3 a second class of matrices in the indefinite inner product space is considered, namely the class of H-expansive matrices. These are matrices A for which $A^*HA - H$ is nonnegative. As in the previous chapter, our purpose is the construction of complex (as well as real) A-invariant maximal H-nonnegative and A-invariant maximal H-nonpositive subspaces. Alternatively, we could have used a suitable Cayley transform to obtain the same results, except for the case when A is real and has both the eigenvalues 1 and -1. The techniques developed in this thesis covers this case. The techniques again entail that a simple form for the matrix pair (A, H) is obtained, with A in Jordan canonical form or in the real canonical form. The simple form is then used to construct the A-invariant maximal H-nonnegative and A-invariant maximal H-nonpositive subspaces.

The last section of the chapter is devoted to obtaining a simple form for the class of H-unitary matrices. We say a real matrix A is H-unitary if $A^THA - H = 0$, for any real symmetric invertible matrix H. What complicates matters here are if A has eigenvalues ± 1 and the corresponding Jordan blocks are of even size.

In the final chapter of the thesis a different topic is considered, namely, rank-one perturbations of H-positive real matrices. A rank-one perturbation of a matrix A is given by $B = A + uv^*$, where u and v are in $\mathbb{C}^n$. We mainly consider perturbations that have the additional structure that uv^* is itself H-positive real. We are interested in the generic behaviour of eigenvalues under such structured rank-one perturbations. A new and interesting result is obtained; that eigenvalues introduced by the perturbation cannot lie on the imaginary axis. However, asymptotically, they may tend towards the imaginary axis.

The chapter ends with a few examples which illustrate an interesting phenomenon which occurs when considering rank-one perturbations in the real case.

Opsomming

Die (bepaalde) inwendige produk van twee vektore $x, y \in \mathbb{R}^n$ word gedefinieer deur $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$. Die lengte van 'n vektor $x \in \mathbb{R}^n$ word dan deur die inwendige produk beskryf as $\|x\| = \langle x, x\rangle^{1/2} = \sqrt{x_1^2 + \ldots + x_n^2}$. In hierdie tesis word die bepaalde inwendige produk vervang deur 'n onbepaalde inwendige produk. Dit het 'n wesenlike impak op die meetkunde van die deelruimtes. Die onbepaalde inwendige produk tussen twee vektore in $\mathbb{R}^n$ word uniek beskryf deur $[x, y] = \langle Hx, y\rangle$, vir 'n reële omkeerbare simmetriese matriks $H = H^T$.

In die literatuur bestaan daar 'n uitgebreide teorie vir die klasse van hermitiese, unitêre en normale matrikse in die bepaalde inwendige produkruimte. Soortgelyke klasse is ook bestudeer in die konteks van die onbepaalde inwendige produkruimte. In hierdie tesis fokus ons op 'n studie van H-positief reële en H-ekspansiewe matrikse. Hierdie studie sluit aan by 'n deeglike studie oor H-dissipatiewe en H-kontraktiewe matrikse, wat reeds in die literatuur bestaan.

In Hoofstuk 2 word 'n klas van matrikse wat nou verband hou met H-dissipatiewe matrikse beskou. Dit is die klas van H-positief reële matrikse. 'n Matriks A is H-dissipatief as $\frac{1}{2i}(HA - A^*H) \geq 0$. 'n $n \times n$ komplekse matriks A is H-positief reëel as $\frac{1}{2}(HA + A^*H) \geq 0$. Dit is maklik om te sien dat A H-positief reëel is as en slegs as iA H-dissipatief is en gevolglik dat $-iA$ H-positief reëel is as en slegs as A H-dissipatief is.

'n Eenvoudige vorm word vir die matriks H verkry en daarna word die A-invariante maksimaal H-nie-negatiewe en A-invariante maksimaal H-nie-positiewe deelruimtes gekonstrueer. Die hoof doelwit is egter om te kyk na die uniekheid en stabiliteit van hierdie deelruimtes. Die numeriese waardeversameling kriterium word hiervoor gebruik. Beide die reële en komplekse gevalle word beskou.

In Hoofstuk 3 word 'n tweede klas van matrikse in die onbepaalde produkruimte beskou, naamlik die klas van H-ekspansiewe matrikse. Dit is matrikse A waarvoor $A^*HA - H$ nie-negatief is. Soos in die vorige hoofstuk is ons doel weereens die konstruksie van komplekse asook reële A-invariante maksimaal H-nie-negatiewe en A-invariante maksimaal H-nie-positiewe deelruimtes. As alternatief tot die metode wat ons gebruik sou dieselfde resultate verkry kon word deur die gebruik van 'n geskikte Cayley transformasie, behalwe vir die geval as A reëel is en beide 1 en -1 as eiewaardes het. Die tegnieke sluit hierdie geval in. Die tegnieke berus weer daarop dat 'n eenvoudige vorm vir die matrikspaar (A, H) verkry word, waarby A in Jordan kanonieke vorm of in reële kanonieke vorm is en dan deur middel van hierdie eenvoudige vorme, soos in Hoofstuk 2, die A-invariante maksimaal H-nie-negatiewe en A-invariante maksimaal H-nie-positiewe deelruimtes te konstrueer.

Die laaste afdeling van die hoofstuk word gewy aan die verkryging van 'n eenvoudige

vorm vir die klas van H-unitêre matrikse. Ons sê dat 'n matriks A H-unitêr is, as $A^T HA - H = 0$ vir enige reële simmetriese omkeerbare matriks H. Wat sake hier bemoeilik is die geval as A eiewaardes ± 1 het en die bybehorende Jordan blokke van ewe grootte is.

Hoofstuk 4 van hierdie proefskrif word aan 'n ander onderwerp gewy, naamlik, rang-een versteurings van H-positief reële matrikse. 'n Rang-een versteuring van 'n matriks A word voorgestel deur $B = A + uv^*$, waar u en v beide in $\mathbb{C}^n$ is. In hierdie hoofstuk word hoofsaaklik versteurings beskou wat sodanig gestruktureer is dat uv^* self ook H-positief reëel is. Ons is geïnteresseerd in die generiese gedrag van die eiewaardes onder hierdie gestruktureerde rang-een versteurings. 'n Nuwe en interessante resultaat word verkry, naamlik dat eiewaardes wat verkry word deur sodanige versteurings nie op die imaginêre as lê nie, maar wel asimptoties na die imaginêre as kan neig. Die hoofstuk eindig met 'n paar voorbeelde wat dié interessante verskynsel in die reële geval van rang-een versteurings illustreer.

Contents

Voorwoord **i**

Summary **ii**

Opsomming **iv**

1 Introduction **1**
1.1 The thesis in more detail 2

2 Uniqueness and stability of H-positive real matrices **19**
2.1 Introduction 19
2.2 Jordan chains and nondegeneracy 19
2.3 The construction of maximal invariant subspaces 27
2.4 The numerical range condition 32
2.5 Uniqueness and stability 38

3 Simple forms and maximal invariant subspaces of H-expansive matrices **46**
3.1 Introductory remarks 46
3.2 Preliminary lemmas 47
3.3 Structure of H-expansive matrices 54
3.4 The general case 80
3.5 Construction of invariant subspaces 82
3.6 On the structure of H-unitary matrices 84

4 Rank-one perturbations **96**
4.1 Introduction 96
4.2 Preliminaries 96
4.3 Main result 99
4.4 Asymptotics of eigenvalues 114
4.5 Rank-one perturbations of H-positive real matrices - The real case 120
4.6 Appendix 122

Bibliography **124**

Index **130**

Chapter 1

Introduction

This thesis is concerned with positive real and expansive matrices in finite dimensional spaces which are equipped with an indefinite inner product. Such matrices occur naturally in several problems in physics and engineering, see e.g. [73, 47] and the references given there. Of particular interest are invariant subspaces of such matrices, which have additional properties with respect to the indefinite inner product. The research area originates with research by Krein and Langer [37]. The study of matrices and operators in an indefinite inner product space is a rich area, see e.g. the books [5, 12, 35, 24, 25]. There are many applications to problems in physics (see, e.g. [22]), factorization theory for rational matrix functions, and systems and control theory (for the last two applications, see, e.g. [7] and [20].)

The theory for positive real and dissipative matrices in an indefinite inner product space and their invariant subspaces has been developed for the case of complex matrices and complex vector spaces, see e.g., [5] for an existence result, and [64, 71, 19] for a constructive approach. It is the aim of the present thesis to extend this theory to the case of real matrices in real indefinite inner product spaces.

In addition, attention will be given in this thesis to perturbation theory for positive real matrices in an indefinite inner product space. Two topics in this area will be discussed in particular. The first topic is that of stability of maximal semidefinite invariant subspaces of such matrices under small perturbations of the matrix. This topic is a natural continuation of a research line started in the early eighties of the last century, see e.g., [26, 6] for the unstructured case. The case of selfadjoint matrices in an indefinite inner product space was treated in [59, 60]; the analogous real cases were treated in [61, 62]. Other structured cases with respect to the indefinite inner product were handled in [63, 52].

The second topic we discuss is the effect of low rank perturbations on the eigenvalues of the matrix. For unstructured matrices this theory has been developed in the recent decades, see [33, 14, 68, 69]. Rank-one perturbations for matrices with certain symmetries in indefinite inner product spaces where first discussed [53] and [54]. This line of research will be continued in this thesis by considering low rank perturbations of matrices which are positive real in indefinite inner product spaces.

In the remainder of this introductory chapter we shall outline in more detail the contents of the thesis, giving the necessary background and definitions, and stating the main results.

1.1 The thesis in more detail

Jordan form

Let us start by recalling the Jordan canonical form of a matrix, both in the complex and in the real case, as this will play an important role in the thesis.

First we consider the vector space $\mathbb{C}^{n\times n}$. All $n \times n$ matrices in this vector space can be subdivided into disjoint equivalence classes of similar matrices, where each equivalence class is determined by a unique matrix canonical form.

We define a typical Jordan block as

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{m\times m}.$$

The fundamental theorem that specifies the Jordan canonical form can for instance be found in [24]. We take the statement of the theorem from there.

Theorem 1.1.1 *For any $A \in \mathbb{C}^{n\times n}$ there is a block-diagonal matrix $J \in \mathbb{C}^{n\times n}$ of the form*

$$J = \begin{bmatrix} J_{m_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & J_{m_r}(\lambda_r) \end{bmatrix},$$

which is similar to A. Moreover, the matrix J is uniquely determined by A up to permutation of the diagonal blocks, i.e., if A is also similar to a matrix

$$\begin{bmatrix} J_{n_1}(\mu_1) & 0 & \dots & 0 \\ 0 & J_{n_2}(\mu_2) & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & J_{n_s}(\mu_s) \end{bmatrix},$$

then $r = s$ and the collection $\{J_{n_1}(\mu_1), J_{n_2}(\mu_2), \dots, J_{n_s}(\mu_s)\}$ (possibly with repeated elements) can be re-arranged so that $J_{n_k}(\mu_k) = J_{m_k}(\lambda_k)$ for $k = 1, 2, \dots, r$.

Proofs of this theorem can be found in many textbooks on linear algebra and on the theory of matrices, e.g. Gantmacher [21], Horn and Johnson [34], Lancaster and Tismenetsky [43].

Secondly we consider matrices over the vector space $\mathbb{R}^{n\times n}$. In the real case it can happen that the eigenvalues are complex, in which case they appear in complex conjugate pairs. Thus, for real numbers, λ and $\mu \neq 0$ we define the real Jordan block of even size,

say $2m \times 2m$ by

$$J_{2m}(\lambda \pm i\mu) = \begin{bmatrix} \lambda & \mu & 1 & 0 & \dots & 0 & 0 \\ -\mu & \lambda & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \mu & \dots & 0 & 0 \\ 0 & 0 & -\mu & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & \lambda & \mu \\ 0 & 0 & 0 & 0 & \dots & -\mu & \lambda \end{bmatrix}.$$

In the above matrix, there are m real 2×2 blocks on the main diagonal and 2×2 identity matrices making up the super-diagonal blocks. Note that $\sigma(J_{2m}(\lambda \pm i\mu)) = \{\lambda + i\mu, \lambda - i\mu\}$, where $\sigma(A)$ denotes the spectrum of the matrix A.

As in the complex case, for any real matrix $A \in \mathbb{R}^{n\times n}$ there exists an invertible $S \in \mathbb{R}^{n\times n}$ such that $J = SAS^{-1}$ is of the form

$$J = J_{m_1}(\lambda_1) \oplus J_{m_2}(\lambda_2) \oplus \dots \oplus J_{m_r}(\lambda_r) \oplus J_{2m_{r+1}}(\lambda_{r+1} \pm i\mu_{r+1}) \oplus \dots \oplus J_{2m_q}(\lambda_q \pm i\mu_q) \quad (1.1)$$

where the λ_j are real and the μ_j are real and positive. The matrix J of (1.1) is uniquely determined by A up to permutation of the diagonal blocks. See, e.g., [24].

The matrix J of (1.1) is known as the real Jordan form of A. A complete proof of this theorem can be found in Gohberg et al. [26]; another source is Shilov [74].

Indefinite inner product

Let $x = (x_1, x_2, \dots, x_n)^T$ be a column vector in $\mathbb{C}^n$. The standard inner product in $\mathbb{C}^n$ is denoted by $\langle ., . \rangle$. Thus,

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_j}$$

where $x = (x_1, x_2, \dots, x_n)^T$, $y = (y_1, y_2, \dots, y_n)^T$ and the bar denotes complex conjugation. As much of the thesis will concern the real case, we emphasize that the same notation is used in $\mathbb{R}^n$.

A function $[., .]$ from $\mathbb{C}^n \times \mathbb{C}^n$ to $\mathbb{C}$ is called an *indefinite inner product* in $\mathbb{C}^n$ if the following axioms are satisfied:

(i) Linearity in the first argument, i.e. $[\alpha x_1 + \beta x_2, y] = \alpha[x_1, y] + \beta[x_2, y]$;

(ii) antisymmetry, i.e. $[x, y] = \overline{[y, x]}$ for all $x, y \in \mathbb{C}^n$;

(iii) nondegeneracy, i.e. if $[x, y] = 0$ for all $y \in \mathbb{C}^n$, then $x = 0$.

For the real case it is noted that the antisymmetry in condition (ii) becomes symmetry. Thus, the function $[., .]$ satisfies all the properties of a standard inner product with the

possible exception that $[x, x]$ may be nonpositive for $x \neq 0$. If H is an invertible Hermitian matrix, then

$$[x, y] = \langle Hx, y\rangle, \quad x, y \in \mathbb{C}^n \tag{1.2}$$

defines an indefinite inner product on $\mathbb{C}^n$. Conversely, for every indefinite inner product $[.,.]$ on $\mathbb{C}^n$, there exists an $n \times n$ invertible and Hermitian matrix H such that (1.2) holds. The same holds in $\mathbb{R}^n$: every indefinite inner product in $\mathbb{R}^n$ is of the form (1.2) for some $n \times n$ invertible symmetric real matrix H.

H-Selfadjoint matrices

For a $n \times n$ complex pair of matrices (A, H) with an indefinite inner product on $\mathbb{C}^n$, we denote the H-adjoint of A as the unique $n\times n$ matrix $A^{[*]}$, which satisfies $[Ax, y] = [x, A^{[*]}y]$ for all $x, y \in \mathbb{C}^n$. The matrix A is said to be H-selfadjoint (or selfadjoint with respect to $[.,.]$) if $A = A^{[*]}$ or, if

$$HA^{[*]} = A^*H.$$

Thus, any H-selfadjoint matrix A is similar to A^*. The converse is also true, that is, if a matrix A is similar to its adjoint then A is H-selfadjoint for some H. In other words, the similarity between A and A^* can be carried out by means of an invertible Hermitian matrix.

For a pair of real $n \times n$ matrices (A, H), such that A is H-selfadjoint, there exists a canonical form for the pair, as stated in the following theorem, from [24]. In the theorem below the sip matrices are matrices of the following form:

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Theorem 1.1.2 *Let (A, H) be a pair of real matrices, where $H = H^T$ is real and invertible, and let A be H-selfadjoint. Then there is an invertible matrix $S \in \mathbb{R}^{n\times n}$ such that $A = S^{-1}JS$ and $H = S^T P_{\epsilon,J} S$, where*

$$J = J_1 \oplus \ldots \oplus J_t \oplus J_{t+1} \oplus \ldots \oplus J_{t+s}$$

is the real Jordan form of A, and

$$P_{\epsilon,J} = \epsilon_1 P_1 \oplus \ldots \oplus \epsilon_t P_t \oplus P_{t+1} \oplus \ldots \oplus P_{t+s}$$

where $P_1, P_2, \ldots, P_{t+s}$ are sip matrices with sizes of $J_1, J_2, \ldots, J_{t+s}$ respectively, and $\epsilon = \{\epsilon_1, \ldots, \epsilon_t\}$ is an ordered set of signs $\epsilon_j = \pm 1, \quad j = 1, \ldots, t$. The set of signs ϵ_j are uniquely determined by (A, H) up to permutation of signs in the blocks of $P_{\epsilon,J}$, corresponding to the Jordan blocks of J with the same real eigenvalue and the same size.

Similar results hold for complex matrix pairs.

H-unitary matrices

A matrix A is said to be H-unitary if it is invertible and $A^{-1} = A^{[*]}$ or $A^{[*]}A = I$, i.e., if $A^*HA = H$.

Canonical forms for H-unitary matrices seem to be less familiar. For the case of Hermitian H, canonical forms have been developed in [28] and also earlier in [25]. Literature can also be found for other special cases of H.

There is however a strong connection between H-unitary and H-selfadjoint matrices. One way to describe this connection is via Cayley transforms. Recall that if $|\alpha| = 1$ and $w \neq \overline{w}$ then the map f defined by

$$f(z) = \alpha(z - \bar{w})/(z - w) \tag{1.3}$$

maps the real line in the z-plane onto the unit circle in the ς-plane, where $\varsigma = f(z)$. The inverse transformation is

$$z = (w\varsigma - \bar{w}\alpha)/(\varsigma - \alpha). \tag{1.4}$$

If $w \notin \sigma(A)$, then the function f is defined on $\sigma(A)$ and, if A is H-selfadjoint, one anticipates that $U = f(A)$ is H-unitary. We rephrase here Proposition 4.3.4 of [24].

Proposition 1.1.3 *Let A be an H-selfadjoint matrix. Let w be a nonreal complex number with $w \notin \sigma(A)$ and let α be any unimodular complex number. Then*

$$U = \alpha(A - \bar{w}I)(A - wI)^{-1} \tag{1.5}$$

is H-unitary and $\alpha \notin \sigma(U)$.

Conversely, if U is H-unitary, $|\alpha| = 1$ and $\alpha \notin \sigma(U)$, then for any $w \neq \bar{w}$ the matrix

$$A = (wU - \bar{w}\alpha I)(U - \alpha I)^{-1} \tag{1.6}$$

is H-selfadjoint and $w \notin \sigma(A)$. Furthermore, the mappings defined by formulas (1.5) and (1.6) are inverses to one another.

The transformation in (1.6) is called the Cayley transform.

H-normal matrices

H-selfadjoint and H-unitary matrices are special cases of H-normal matrices. A matrix $A \in \mathbb{C}^{n\times n}$ is called H-normal if A commutes with its H-adjoint, i.e., if $AA^{[*]} = A^{[*]}A$. (See, [48].)

Finding a canonical form for H-normal matrices has proven to be as difficult as classifying pairs of commuting matrices under simultaneous similarity, see [29]. So far, a classification of H-normal matrices has only been obtained for some special cases, see [29, 30, 31]. In the paper [30], Theorem 2, the classification of indecomposable H-normal operators can be found.

It therefore makes sense to look for proper subsets for which a complete classification can be obtained. A first approach has been made in [28], where block-Toeplitz H-normal

matrices have been defined. Thereafter, a complete classification for block-Toeplitz H-normal matrices was given in [27] for the case that H induces a Hermitian sesquilinear form. However, in the case that H induces a complex or real bilinear form that is symmetric or skew-symmetric, there exist H-selfadjoint, H-skewadjoint, or H-unitary matrices that fail to be block-Toeplitz H-normal. Thus, the approach via block-Toeplitz H-normal matrices only makes sense for the case of a Hermitian sesquilinear form.

In [56], several subsets of the set of H-normal matrices have been considered with the emphasis on finding a subset that is 'large enough' in order to contain all H-selfadjoint, H-skewadjoint, and H-unitary matrices, but that is still small enough such that a complete classification of its elements can be obtained. In [48] a suitable set with these properties is described: namely, the set of polynomially H-normal matrices. By definition, a matrix $X \in \mathbb{C}^{n\times n}$ is polynomially H-normal if there exists a polynomial $p \in \mathbb{C}[t]$ such that $X^{[*]} = p(X)$.

Normal operators on infinite dimensional indefinite inner product spaces have been studied in [9], [10] and [82].

H-dissipative matrices

Let $H = H^*$ be an invertible $n \times n$ matrix, then an $n \times n$ matrix A is called *H-dissipative* if $\frac{1}{2i}(HA - A^*H) \geq 0$, or equivalently, for any vector $x \in \mathbb{C}^n$, $\mathrm{Im}\langle HAx, x\rangle \geq 0$.

In [64] it is shown how a matrix pair (A, H), where $H = H^*$ is invertible and A is H-dissipative, can be reduced to a simple form. That is, an invertible matrix S is found such that $S^{-1}AS$ is in Jordan canonical form and S^*HS has a form which is as simple as needed by the context.

We start with A in Jordan canonical form. The simple form for the matrix H is in the following corollary, (Corollary 2.7, p 179 of [64].)

Corollary 1.1.4 *Let $A = diag[J_1, \ldots, J_N]$ with J_k an $n_k \times n_k$ Jordan block with eigenvalue λ_k for $k = 1, \ldots N$. Let $H = (H_{k,l})_{k,l=1}^N$ be an invertible Hermitian matrix in corresponding block form. If A is H-dissipative, then for $k, l = 1, \ldots N$,*

1. *if $\lambda_k \in \mathbb{R}$, then $H_{k,k}$ has the form*

$$H_{k,k} = \left[\begin{array}{ccc|ccc} & & & & & a \\ & 0 & & & \iddots & \\ & & a & & & \iddots \\ \hline & \iddots & & & * & \\ a & & \iddots & & & \end{array}\right]$$

when n_k is odd, and

$$H_{k,k} = \left[\begin{array}{cc|cc} & & & \bar{a} \\ & 0 & & \iddots \\ & & \bar{a} & \iddots \\ \hline & a & & \\ & \iddots & & * \\ a & \iddots & & \end{array}\right],$$

when n_k is even. The lines in the matrices divide the matrices into blocks such that the left upper block has size $m_k \times m_k$, with $m_k = \frac{1}{2}(n_k + 1)$ in the odd case and $m_k = \frac{1}{2}n_k$ in the even case. In the left upper triangle the entries are 0. Starting with the skew diagonal, a Hankel structure appears in the right upper and the left lower block. In the right lower block there is no such particular pattern ;

2. *if $\lambda_k = \lambda_l \in \mathbb{R}$ then $H_{k,l}$ has the form*

$$H_{k,l} = \left[\begin{array}{cc|cc} & & & \iddots \\ & 0 & 0 & \iddots \; \iddots \\ & & \iddots & \iddots \\ & & \iddots \; \iddots & \iddots \\ \hline & \iddots \; \iddots & & \\ & 0 \quad \iddots \; \iddots & & \\ & \iddots \; \iddots & & * \\ \iddots & \iddots & & \end{array}\right]$$

when $n_k < n_l$, and the transposed structure when $n_k > n_l$;

$$H_{k,l} = \left[\begin{array}{cc|cc} & & & \iddots \\ & 0 & 0 & \iddots \; \iddots \\ & & \iddots & \iddots \\ & & \iddots \; \iddots & \iddots \\ \hline & \iddots & & \\ & 0 \quad \iddots \; \iddots & & \\ & \iddots \; \iddots & & * \\ \iddots & \iddots & & \end{array}\right]$$

when $n_k = n_l$ is even;

$$H_{k,l} = \left[\begin{array}{ccc|ccc} & 0 & & & 0 & \\ & & & & & \\ & & & & & \\ \hline & 0 & & & * & \\ & & & & & \\ & & & & & \end{array}\right]$$

when $n_k = n_l$ is odd. The horizontal and vertical lines divide the matrices $H_{k,l}$ into blocks such that the left upper block is of size $m_k \times m_l$. In case $n_k = n_l$ is odd, there is only one nonzero entry in the left upper block. In all cases the off-diagonal blocks of $H_{k,l}$ have a Hankel structure.

3. *if $\lambda_k, \lambda_l \in \mathbb{R}, \lambda_k \neq \lambda_l$, then $H_{k,l}$ has the form*

$$H_{k,l} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & * \end{array}\right]$$

where the left upper zero block is of size $m_k \times m_l$;

4. *if $\lambda_k \in \mathbb{R}, \lambda_l \in \mathbb{C} \setminus \mathbb{R}$, then $H_{k,l}$ has the form*

$$H_{k,l} = \left[\begin{array}{c} 0 \\ \hline * \end{array}\right]$$

where the upper zero block is of size $m_k \times n_l$.

H-positive real matrices over the complex field and real field

This class of matrices is closely related to H-dissipative matrices. We will say that an $n \times n$ matrix A is H-positive real if $\frac{1}{2}(HA + A^*H) \geq 0$. It is easily seen that A is H-positive real if and only if iA is H-dissipative, and hence if follows that $-iA$ is H-positive real if and only if A is H-dissipative.

The paper [19] considers both the complex and real case. In the real case, $H = H^T$ is an invertible real symmetric matrix, and A is a real matrix satisfying $HA + A^TH \geq 0$, i.e. A is H-positive real. Compared to the complex case the additional difficulty is that the eigenvalues now appear in complex conjugate pairs, and that we shall be interested in real A-invariant maximal H-nonnegative subspaces. The real Jordan form is used instead of the complex Jordan form. Recall that the real Jordan form of a real matrix A is given by $A = SJS^{-1}$, where S is a real invertible matrix and $J = \mathrm{diag}(J_1, \ldots, J_N)$, where J_i is either a standard Jordan block with a real eigenvalue, or a so-called real Jordan block corresponding to a pair of complex conjugate eigenvalues $a \pm bi$, given by

$$J_i = \begin{bmatrix} \alpha & I & & \\ & \ddots & \ddots & \\ & & \ddots & I \\ & & & \alpha \end{bmatrix},$$

where $\alpha = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

The simple form for the matrix H in the case where $\sigma(A)$ is just a pair of complex conjugate pure imaginary eigenvalues is given in the next theorem. The 2×2 matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is denoted by H_0.

Theorem 1.1.5 *Let $A = J_1 \oplus J_2 \oplus \dots J_N$ be H-positive real, where each J_i is a $2n_i \times 2n_i$ matrix of the form*

$$A = \begin{bmatrix} \alpha & I & & \\ & \ddots & \ddots & \\ & & \ddots & I \\ & & & \alpha \end{bmatrix},$$

where $\alpha = aH_0$ and $a \neq 0$. Then $H = [H^{(i,j)}]_{i,j=1}^{N}$, where $H^{(i,j)}$ is an $2n_i \times 2n_j$ matrix of the form

$$[H^{(i,j)}] = (H_{kl}^{(i,j)})_{k=1,l=1}^{n_i,n_j},$$

and each $H_{kl}^{(i,j)}$ is a 2×2 matrix. Moreover, for $i = j$ and $n = n_i$ is even *H has the form*

$$H^{(i,i)} = \left[\begin{array}{ccc|ccc} & & & & & H_{1n}^{(i,i)} \\ & & & & & -H_{1n}^{(i,i)} \\ & 0 & & & \iddots & \\ & & & & (-1)^{\frac{n-2}{2}} H_{1n}^{(i,i)} & \iddots \\ \hline & & (-1)^{\frac{n-2}{2}} H_{1n}^{(i,i)T} & & & \\ & \iddots & & & * & \\ & -H_{1n}^{(i,i)T} & & & & \\ H_{1n}^{(i,i)T} & & \iddots & & & \end{array}\right]$$

where $H_{kl}^{(i,i)}$ is of the form $a_{kl}^{(i,i)} I + b_{kl}^{(i,i)} H_0$ for some $a_{kl}^{(i,i)}, b_{kl}^{(i,i)} \in \mathbb{R}$, in case $\max(k,l) \leq \frac{n}{2}$ and if $n = n_i$ is odd

$$H^{(i,i)} = \left[\begin{array}{ccc|ccc} & & & & & H_{1n}^{(i,i)} \\ & & & & -H_{1n}^{(i,i)} & \\ & 0 & & \iddots & & \\ & & & (-1)^{\frac{n-3}{2}} H_{1n}^{(i,i)} & & \\ & & (-1)^{\frac{n-1}{2}} H_{1n}^{(i,i)} & & & \iddots \\ \hline & & (-1)^{\frac{n-3}{2}} H_{1n}^{(i,i)T} & & & \\ & \iddots & & & * & \\ & -H_{1n}^{(i,i)T} & & & & \\ H_{1n}^{(i,i)T} & & \iddots & & & \end{array}\right]$$

where $H_{kl}^{(i,i)}$ is of the form $a_{kl}^{(i,i)} I + b_{kl}^{(i,i)} H_0$ for some $a_{kl}^{(i,i)}, b_{kl}^{(i,i)} \in \mathbb{R}$ in case $\max(k,l) \leq \frac{n+1}{2}$ and $H_{1n}^{(i,i)} = a_{1n}^{(i,i)} I$.

The off-diagonal blocks in H are real lower triangular alternating block 2×2 Hankel matrices, i.e. it is of the form

$$H^{(i,j)} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & (-1)^{n-1}H_1^{(i,j)} \\ \vdots & & & & & & & 0 & (-1)^{n-2}H_1^{(i,j)} & (-1)^{n-2}H_2^{(i,j)} \\ \vdots & & & & & & \iddots & \iddots & & \vdots \\ \vdots & & & & & \iddots & -H_1^{(i,j)} & -H_2^{(i,j)} & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & H_1^{(i,j)} & H_2^{(i,j)} & \cdots & \cdots & H_n^{(i,j)} \end{bmatrix},$$

with

$$H_k^{(i,j)} = \begin{bmatrix} a_k^{(i,j)} & b_k^{(i,j)} \\ -b_k^{(i,j)} & a_k^{(i,j)} \end{bmatrix}, \qquad a_k^{(i,j)}, b_k^{(i,j)} \in \mathbb{R}.$$

The fact that H is symmetric forces the relation between the block in the upper right hand corner and the block in the lower left hand corner, as well as the specific form of the block on the main anti-diagonal in the case when n is odd.

In case A is a single Jordan block with zero eigenvalue then the form of H is as in Corollary 1.1.4.

Something more can be observed from the proof of Theorem 3.1 in [19]. The observation to follow here, helps in understanding how the eigenvalues of a rank-one perturbation move away from the imaginary axis.

First consider the following example.

Example 1.1.6

Let $A = \begin{bmatrix} \gamma H_0 & I_2 \\ 0 & \gamma H_0 \end{bmatrix}$, and $H = \begin{bmatrix} 0 & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix}$, where $H_{12} = aI_2 + bH_0$, for some $a, b \in \mathbb{R}$. Assume that $HA + A^T H \geq 0$. A direct computation gives

$$HA + A^T H = \begin{bmatrix} 0 & 0 \\ 0 & H_{12} + H_{12}^T + \gamma H_{22} H_0 - \gamma H_0 H_{22} \end{bmatrix}.$$

Now $H_{22} = \begin{bmatrix} c & d \\ d & e \end{bmatrix}$ for some real numbers c, d and e. Further, $H_{12} + H_{12}^T = 2aI_2$. Computing the $(2,2)$ block entry, which we denote by D, gives

$$D = \begin{bmatrix} 2a - 2\gamma d & \gamma(c-e) \\ \gamma(c-e) & 2a + 2\gamma d \end{bmatrix}.$$

Now for A to be H-positive real the matrix D needs to positive semidefinite. In particular, both diagonal entries of D need to be greater than or equal to zero. So $2a \geq 2\gamma d$ and $2a \geq -2\gamma d$. It follows that $2a \geq |2\gamma d| \geq 0$. In particular we conclude that $a \geq 0$. Something similar holds for all blocks of even size with pure imaginary eigenvalues.

So, let us continue the arguments from the proof of Theorem 3.1 in [19] (or Theorem 1.1.5 stated above). In particular, we follow this up starting from formula (29) in [19]. We assume, using the notation there, that n is even, say $n = 2k$, and we consider the

$(k+1,k+1)$-block-entry of $HA+A^TH$. Since A is H-positive real, this two by two matrix must be positive semidefinite. According to formula (29) it is given by

$$(H_{k+1,k+1}\alpha - \alpha H_{k+1,k+1}) + H_{k+1,k} + H_{k,k+1}.$$

Using Theorem 1.1.5 we have $H_{k,k+1} = (-1)^{k-1}H_{1n}$, and $H_{k+1,k} = (-1)^{k-1}H_{1n}^T$. Moreover, we have $H_{1n} = a_{1n}I + b_{1n}H_0$. The symmetric two by two matrix $H_{k+1,k+1}$ we denote by $H_{k+1,k+1} = \begin{bmatrix} c & d \\ d & e \end{bmatrix}$. Then a straightforward computation yields

$$(H_{k+1,k+1}\alpha - \alpha H_{k+1,k+1}) = a \begin{bmatrix} -2d & c-e \\ c-e & 2d \end{bmatrix},$$

while $H_{k+1,k} + H_{k,k+1} = (-1)^{k-1}2a_{1n}I_2$. Thus

$$(H_{k+1,k+1}\alpha - \alpha H_{k+1,k+1}) + H_{k+1,k} + H_{k,k+1} = \begin{bmatrix} (-1)^{k-1}2a_{1n} - 2da & a(c-e) \\ a(c-e) & (-1)^{k-1}2a_{1n} + 2da \end{bmatrix}.$$

This matrix being positive semidefinite implies in particular that its diagonal entries are greater than or equal to zero. In particular we see that $(-1)^{k-1}2a_{1n} \geq |2da| \geq 0$. Thus we conclude that for even n the value of a_{1n} in the block H_{1n} satisfies $(-1)^{\frac{n-2}{2}}a_{1n} \geq 0$.

The simple forms for H in general, are given in Theorem 3.9 of [19] as follows:

Theorem 1.1.7 *Let $A = diag(J_1, \ldots, J_N)$, where either J_k is an $n_k \times n_k$ Jordan block with real eigenvalue λ_k, or J_k is a $2n_k \times 2n_k$ real Jordan block with a pair of complex conjugate eigenvalues $\lambda_k, \bar{\lambda}_k$, for $k = 1, \ldots, N$. Define $m_k = [(n_k+1)/2]$. Let $H = (H_{k,l})_{k,l=1}^N$ be a real invertible symmetric matrix in corresponding block form. If A is H-positive real, then for $k, l = 1, \ldots, N$ we have*

(a1) *if $\lambda_k = 0$ then $H_{k,k}$ is as in Corollary 1.1.4,*

(a2) *if J_k has a pair of complex conjugate pure imaginary nonzero eigenvalues, say $a_k i$ and $-a_k i$, then $H_{k,k}$ is as in Theorem 1.1.5,*

(b1) *if J_k and J_l both have zero eigenvalue, then*

$$H_{k,l} = \left[\begin{array}{c|c} 0 & H_1 \\ \hline H_2 & * \end{array}\right],$$

with H_1 and H_2 lower triangular alternating Hankel matrices.

(b2) *if J_k and J_l both have pure imaginary nonzero eigenvalue $\pm a_k i$ and $\pm a_l i$ with $a_k = a_l$ then*

$$H_{k,l} = \left[\begin{array}{cc|c} 0 & & \\ & a & H_1 \\ \hline & H_2 & * \end{array}\right],$$

where H_1 and H_2 are (square) lower triangular alternating Hankel matrices and a is some complex number, which depends on k and l.

(c1) *if J_k has zero eigenvalue, and J_l has a pair of complex conjugate pure imaginary nonzero eigenvalues $\pm a_l i$, then $H_{k,l}$ has the form*

$$H_{k,l} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & * \end{array}\right], \tag{1.7}$$

where the left upper zero block is of size $m_k \times 2m_l$,

(c2) *if J_k and J_l both have a pair of complex conjugate pure imaginary nonzero eigenvalues $\pm a_k i$, respectively $\pm a_l i$, then $H_{k,l}$ has the form* (1.7) *where now the left upper zero block is of size $2m_k \times 2m_l$,*

(d1) *if J_k has zero eigenvalue, and J_l has a real nonzero eigenvalue λ_l, then $H_{k,l}$ has the form*

$$H_{k,l} = \left[\begin{array}{c} 0 \\ \hline * \end{array}\right], \tag{1.8}$$

where the upper zero block is of size $m_k \times n_l$,

(d2) *if J_k has zero eigenvalue, and J_l has a pair of complex conjugate eigenvalues $a_l \pm b_l i$, with $a_l, b_l \neq 0$ then $H_{k,l}$ has the form* (1.8) *where now the upper zero block is of size $m_k \times 2n_l$,*

(d3) *if J_k has a pair of complex conjugate pure imaginary nonzero eigenvalues $\pm a_k i$ and J_l has a real nonzero eigenvalue λ_l then $H_{k,l}$ has the form* (1.8) *where now the upper zero block is of size $2m_k \times n_l$,*

(d4) *if J_k has a pair of complex conjugate pure imaginary nonzero eigenvalues $\pm a_k i$ and J_l has a pair of complex conjugate eigenvalues $a_l \pm b_l i$, with $a_l, b_l \neq 0$ then $H_{k,l}$ has the form* (1.8) *where now the upper zero block is of size $2m_k \times 2n_l$.*

In Chapter 2 of this thesis we continue the investigation of invariant maximal nonnegative subspaces for H-positive real matrices, which started in [19]. There, an explicit construction was given for the existence of such subspaces, based on the simple form above. For the readers' convenience, this construction is largely repeated in Chapter 2, and refined. In particular, an analoque of the numerical range condition, introduced in [77] for H-dissipative matrices, is considered here. The uniqueness and stability of A-invariant maximal H-nonnegative subspaces is the main focus of the chapter. In Theorem 5.3 the equivalence is proved between uniqueness, stability and the numerical range condition. The real case is treated largely as a consequence of the complex case.

H-expansive matrices

Let $H = H^T$ be an invertible $n \times n$ symmetric real matrix, and let A be an $n \times n$ real matrix satisfying $A^T H A \geq H$ or $A^T H A - H \geq 0$. Such a matrix is called H-expansive. This is equivalent to $[Ax, Ax] \geq [x, x]$ for all $x \in \mathbb{R}^n$.

We have already seen that there is a connection through Cayley transform between H-unitary and H-selfadjoint matrices. This is also true for complex H-expansive and complex H-positive real matrices.

Let $H = H^*$ be an invertible $n \times n$ Hermitian matrix, and let A be an $n \times n$ complex matrix satisfying $A^*HA \geq H$. Again, such a matrix is called H-expansive. Let

$$B = (\eta + A)(\eta - A)^{-1},$$

with $\eta \notin \sigma(A)$ and $|\eta| = 1$. Note that $HB + B^*H = 2((\eta - A)^*)^{-1}(H - A^*HA)(\eta - A)^{-1} \leq 0$, that is $-B$ is H-positive real. (It can be somewhat confusing that this property is called H-dissipative in [24], see pages 61, 218 and following.) If $\eta \notin \mathbb{R}$, then B will be a complex matrix, even if A and H were real matrices. Therefore the theory with respect to complex matrices can be applied, see [19]. Take note that A can also be expressed in terms of B, that is

$$A = (B + I)^{-1}(B - I)\eta,$$

since $\eta \notin \sigma(A)$ implies $-1 \notin \sigma(B)$.

For the real case, if $1 \notin \sigma(A)$ or $-1 \notin \sigma(A)$ then the same argument gives

$$B = (I + A)(I - A)^{-1} \quad \text{or} \quad B = (-I + A)(-I - A)^{-1},$$

respectively. Thus, if A is real so is B. Again, in this case $HB + B^TH \leq 0$. However, if both 1 and -1 are in $\sigma(A)$, then we cannot directly apply this argument. We therefore consider the real case without using the Cayley transform.

In Chapter 3 of the thesis we develop for a real symmetric invertible H and for a real H-expansive matrix A a simple form which is comparable to the one developed in [19] for real H-positive real matrices. This simple form may then be used in the construction of A-invariant maximal H-nonnegative and A-invariant maximal H-nonpositive subspaces. In addition, in that chapter we make some observations concerning real H-unitary matrices. Canonical forms for H-unitary matrices exist, see e.g. [24] and [66]. These canonical forms are largely derived by means of Cayley transform. We conjecture that direct methods such as used in Chapter 2 may lead to a different (obviously equivalent) canonical form. First steps in this direction are taken in the last section of Chapter 3.

H-Hamiltonian (real case)

If $K_n := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$, where I_n is the $n \times n$ identity matrix, we call a matrix $\mathcal{H} \in \mathbb{C}^{2n \times 2n}$ Hamiltonian if $\mathcal{H}K_n = (\mathcal{H}K_n)^*$. Here $*$ represents the conjugate transpose. One can find the canonical form for a real Hamiltonian matrix $\mathcal{H}$ in the paper of Lin et al [46]. (See Theorem 22, page 497.) Other canonical forms for Hamiltonian matrices can also be found in this article.

H-symplectic real

With the notation as in the previous paragraph, we call a matrix $\mathcal{S} \in \mathbb{C}^{2n \times 2n}$ symplectic if $\mathcal{S}^*K_n\mathcal{S} = K_n$. A proof for the canonical form of a real symplectic matrix $\mathcal{S}$ can also be found in [46]. (See Theorem 41, page 529.)

Invariant subspaces

Definition 1.1.8 *A subspace $\mathcal{M} \subseteq \mathbb{C}^n$ is said to be invariant for an $n \times n$ matrix A, or to be A-invariant, if $x \in \mathcal{M}$ implies $Ax \in \mathcal{M}$.*

Definition 1.1.9 *A vector x in $\mathbb{C}^n$ is called H-nonnegative if $[x, x] \geq 0$, H-nonpositive if $[x, x] \leq 0$, and H-neutral if $[x, x] = 0$. In the same way, a subspace $\mathcal{M} \subset \mathbb{C}^n$ is called H-nonnegative, H-nonpositive and H-neutral, if each vector in the subspace is H-nonnegative, H-nonpositive, both H-nonnegative and H-nonpositive, respectively.*

Definition 1.1.10 *A subspace $\mathcal{M} \subseteq \mathbb{C}^n$ is called maximal H-nonnegative if it is not contained in a larger H-nonnegative subspace.*

In Section 5.12, Theorem 5.12.1 of [24], it is shown that every H-selfadjoint matrix has an invariant subspace which is maximal H-nonnegative (or maximal H-nonpositive). The proof was based on the canonical form for the pair (A, H). In an infinite dimensional setting, the result is in fact older, see [58], [76] and [37].

The explicit construction of H-nonnegative and H-nonpositive subspaces when A is H-positive real can be found in Section 3.4 of [19].

Existence of maximal semidefinite invariant subspaces of dissipative operators in an infinite dimensional indefinite inner product space was first shown in the papers of Azizov et al., see [2], [3], [4] and also [5]. In the paper [72] the explicit construction for invariant subspaces of dissipative operators in a space with indefinite metric can be found. In the PhD thesis [77], the construction for A-invariant H-nonnegative and H-nonpositive subspaces, where A is a densely defined closed dissipative operator on the Pontryagin space Π_κ and λ is a real eigenvalue of A, can be found. The stability of the subspaces is discussed in Section 3.3 of [77]. A discussion of the uniqueness of the H-nonnegative and H-nonpositive subspaces, when A is an H-dissipative matrix, can be found in Theorem 2.1 of [63].

The second chapter of the thesis completes a study already started in [19] pertaining to the positive real case. Here, not only the complex case, but also the real case is treated. The following results (in the form of Theorems) are obtained regarding the uniqueness and stability of the A-invariant subspaces, where $A \in \mathbb{A}$ and $\mathbb{A}$ denotes the class of all $n \times n$ H-positive real matrices.

Theorem 1.1.11 *The following statements are equivalent for a given H-positive real matrix A:*

(i) there exists an $\mathbb{A}$-stable A-invariant maximal H-nonnegative subspace, say $\mathcal{M}_+$.

(ii) there exists an $\mathbb{A}$-stable A-invariant maximal H-nonpositive subspace, say $\mathcal{M}_-$.

(iii) the numerical range condition holds for the pair (iA, H).

(iv) there is a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, with $\sigma(A|_{\mathcal{M}_+})$ contained in the closed right half plane.

(v) there is a unique A-invariant maximal H-nonpositive subspace $\mathcal{M}_-$, with $\sigma(A|_{\mathcal{M}_-})$ contained in the closed left half plane.

Lemma 1.1.12 *The following statements are equivalent for a given real H-positive real matrix A.*

(a) There exists a unique complex A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, such that $\sigma(A|_{\mathcal{M}_+}) \subseteq \overline{\mathbb{C}}_+$.

(b) There exists a unique real A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, such that $\sigma(A|_{\mathcal{M}_+}) \subseteq \overline{\mathbb{C}}_+$.

(c) There exists a unique complex A-invariant maximal H-nonpositive subspace $\mathcal{M}_-$, such that $\sigma(A|_{\mathcal{M}_-}) \subseteq \overline{\mathbb{C}}_-$.

(d) There exists a unique real A-invariant maximal H-nonpositive subspace $\mathcal{M}_-$, such that $\sigma(A|_{\mathcal{M}_-}) \subseteq \overline{\mathbb{C}}_-$.

Low rank perturbations

A lot is known about generic rank-one perturbation of matrices. Consider a complex $n \times n$ matrix A and arbitrary vectors $u, v \in \mathbb{C}^n$. We now want to study the eigenvalues of the following matrix as function of t:

$$B(t) = A + tuv^*.$$

The above equation can be motivated by considering a single input control system, namely

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned}$$

where A is $n \times n$, B is $n \times 1$, C is $m \times n$, and D is $m \times 1$. These type of models often appear in engineering problems. The state $x(t)$ usually has coordinates that are position and speed, the input $u(t)$ might be power (determining acceleration), and the outputs $y(t)$ may be measurements of position, speed, etc.

If we consider the input-state equation:

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

its solution is

$$x(t) = e^{At}x(0) + \int_0^t e^{(t-s)A}Bu(s)\,ds.$$

Usually we expect states to go to zero when there is no input (stability). It is well known that we have stability if all eigenvalues of A are in the open left half plane.

Thus, we would like to find a controller (a new system) which stabilizes the system. If we assume for simplicity that the full state is available for control, we try $u(t) = Fx(t)$. Then the closed loop system becomes

$$\dot{x}(t) = Ax(t) + Bu(t) = (A + BF)x(t).$$

So, if we can find a matrix F so that $A+BF$ has all its eigenvalues in the open left half plane we have stability i.e., we have found a stabilizing feedback. Observe that $A + BF$ is a rank-one perturbation of A. We will explain the basics behind perturbation theory by means of a few examples where the perturbations considered are of rank two.

Example 1.1.13

Let $A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, with $\lambda_1, \lambda_2 \in \mathbb{R}$ and let $B = \begin{bmatrix} \alpha & b \\ \overline{b} & \gamma \end{bmatrix}$, with $\alpha, \gamma \in \mathbb{R}$ and $b \in \mathbb{C}$. Now consider the self-adjoint perturbation $A + \epsilon B$, with $\epsilon \in \mathbb{C}$, i.e.

$$A + \epsilon B = \begin{bmatrix} \lambda_1 + \epsilon\alpha & \epsilon b \\ \epsilon\overline{b} & \lambda_2 + \epsilon\gamma \end{bmatrix}.$$

The characteristic polynomial of $A + \epsilon B$ reduces to the following expression:

$$\begin{aligned} p\,(A + \epsilon B) &= \det(\lambda I - (A + \epsilon B)) \\ &= (\lambda - \lambda_1(\epsilon))(\lambda - \lambda_2(\epsilon)), \end{aligned}$$

where

$$\begin{aligned} \lambda_{1,2}(\epsilon) &= \frac{\lambda_1 + \lambda_2 + \epsilon(\alpha + \gamma)}{2} \\ &\pm \left[\frac{(\lambda_1 - \lambda_2)^2}{4} + \frac{\epsilon}{2}(\lambda_1 - \lambda_2)(\alpha - \gamma) + \epsilon^2(|b|^2 + \frac{1}{4}(\alpha - \gamma)^2)\right]^{1/2}. \end{aligned} \tag{1.9}$$

When, for example, we have $\lambda_1 = \lambda_2$ then

$$\begin{aligned} \lambda_{1,2}(\epsilon) &= \lambda_1 + \frac{1}{2}(\alpha + \gamma)\epsilon \pm (|b|^2 + \frac{1}{4}(\alpha - \gamma)^2)^{1/2}\epsilon \\ &= \lambda_1 + [\frac{1}{2}(\alpha + \gamma) \pm (|b|^2 + \frac{1}{4}(\alpha - \gamma)^2)^{1/2}]\epsilon, \end{aligned}$$

which can be seen as a power series in ϵ in a neighbourhood of $\epsilon = 0$. When we set $\epsilon = 0$ in (1.9), we see that we obtain the eigenvalues of A again, as expected. Hence, the behavior of the eigenvalues depend on the magnitude of ϵ. (For more details, see [8], page 99 and [36]).

We now give an example of a rank two perturbation.

Example 1.1.14

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, and let $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Consider

$$A + \epsilon B = \begin{bmatrix} 1 + \epsilon & 0 & 0 \\ 0 & 0 & 1 \\ 0 & \epsilon & 0 \end{bmatrix}.$$

Hence we have a rank two perturbation of the matrix A. The characteristic polynomial of $A + \epsilon B$ reduces to the following:

$$\begin{aligned} p\,(A + \epsilon B) &= \det(\lambda I - (A + \epsilon B)) \\ &= (\lambda - (1 + \epsilon))(\lambda^2 - \epsilon). \end{aligned}$$

Putting the characteristic polynomial equal to zero gives the following eigenvalues

$$\lambda_1 = 1 + \epsilon, \qquad \lambda_2 = \sqrt{\epsilon}, \qquad \lambda_3 = -\sqrt{\epsilon}.$$

Again with $\epsilon = 0$, the original eigenvalues of A are obtained, but if $\epsilon > 0$, the eigenvalues of $A + \epsilon B$ can change quite dramatically. For more on Puisseux series, see [36], see also [57].

The above example can be generalized to generic rank-one perturbations of the Jordan form of matrices, e.g.,

Theorem 1.1.15 *Let $A \in \mathbb{R}^{n\times n}$ be a matrix having the pairwise distinct eigenvalues $\lambda_1, \ldots, \lambda_p$ with geometric multiplicities $g_1, \ldots g_p$ and having the Jordan canonical form*

$$\bigoplus_{k=1}^{g_1} J_{n_1,k}(\lambda_1) \oplus \ldots \oplus \bigoplus_{k=1}^{g_p} J_{n_p,k}(\lambda_p),$$

where $n_{j,1} \geq \ldots \geq n_{j,g_j}$, $j = 1, \ldots, p$. Consider the rank-one matrix $B = uv^T$, with $u, v \in \mathbb{R}^n$. Then generically (with respect to the entries of u and v) the Jordan blocks of $A + B$ with eigenvalue λ_j are just the $g_j - 1$ smallest Jordan blocks of A with eigenvalue λ_j, and all other eigenvalues of $A + B$ are simple; if $g_j = 1$, then generically λ_j is not an eigenvalue of $A + B$.

More precisely, there is a generic set $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^n$ such that for every $(u, v) \in \Omega$, the Jordan structure of $A + uv^T$ is described in (a) and (b) below:

(a) the Jordan structure of $A + uv^T$ for the eigenvalues $\lambda_1, \ldots, \lambda_p$ is given by

$$\bigoplus_{k=2}^{g_1} J_{n_1,k}(\lambda_1) \oplus \ldots \oplus \bigoplus_{k=2}^{g_p} J_{n_p,k}(\lambda_p);$$

(b) the eigenvalues of $A + uv^T$ that are different from any of $\lambda_1, \ldots, \lambda_p$, are simple.

The proof of Part (a) of Theorem 1.1.15 was given in [33], later also in [14, 68]. A similar result for low rank perturbation based on spectral resolvent was obtained in [69]. Numerous papers (see the corresponding references in [36] and [8] are devoted to the construction of the perturbation theory for eigenvalues of $A + tB$ that do not belong to the spectrum of A (here t is the perturbation parameter). An alternative proof of Part (a) of Theorem 1.1.15, using the Brunovsky canonical form, was recently given in [53].

Rank-one perturbations for structured matrices in indefinite inner product spaces was studied in [54], and for the J-Hamiltonian case in [53]. In [55] the effect on the canonical form of Theorem 2.3 is considered. In [53] it is shown that structured rank-one perturbations of J-Hamiltonian matrices leads to a surprising new result. The next example shows this. In the canonical form of real J-Hamiltonian matrices Jordan blocks with eigenvalue zero of odd size must come in pairs.

Example 1.1.16

Let

$$A = \begin{bmatrix} J_{2m+1}(0) & 0 \\ 0 & J_{2m+1}(0) \end{bmatrix}, \quad H = \begin{bmatrix} 0 & \Sigma_{2m+1} \\ \Sigma_{2m+1} & 0 \end{bmatrix},$$

where

$$\Sigma_k = \begin{bmatrix} & & & 1 \\ & & -1 & \\ & \iddots & & \\ (-1)^{k-1} & & & \end{bmatrix}.$$

Then a generic unstructured rank-one perturbation $A + uv^*$ would have, according to Theorem 1.1.15, a Jordan block $J_{2m+1}(0)$ and $2m+1$ simple non-zero eigenvalues. Surprisingly in contrast, a structured rank-one perturbation of the form $A + uu^*H$ has a Jordan block $J_{2m+2}(0)$ and only $2m$ simple non-zero eigenvalues.

In Chapter 4 we consider structured rank-one perturbation of H-positive real matrices A, that is, perturbations of the form $A + tuu^*H$ for generic vectors u and $t > 0$. It turns out that a perturbation uv^* is H-positive real if and only if $v^* = cu^*H$ for some c with $\mathrm{Re}\, c \geq 0$. A new and interesting result is obtained, where the constant c is replaced with the real positive constant t. It is proved that the following holds:

Proposition 1.1.17 *For all vectors u and all $t > 0$, $\sigma(A + tuu^*H) \backslash \sigma(A) \cap i\mathbb{R} = \emptyset$. In other words, eigenvalues introduced by the perturbation cannot lie on the imaginary axis. However, asymptotically as $t \to \infty$ they may tend towards the imaginary axis.*

An analogous result for low rank perturbations is proved as well. In addition we shall show the analogue of Theorem 1.1.15 for this case.

Chapter 2

Uniqueness and stability of H-positive real matrices

2.1 Introduction

In this paper we study a class of matrices that is closely related to *H-dissipative matrices*. We will say an $n \times n$ complex matrix A is *H-positive real* if $\frac{1}{2}(HA + A^*H) \geq 0$. Recall that a complex matrix A is H-dissipative if $\frac{1}{2i}(HA - A^*H) \geq 0$. It is easily seen that A is H-positive real if and only if iA is H-dissipative, and hence it follows that $-iA$ is H-positive real if and only if A is H-dissipative.

Section 2.2 follows the line of argument of [19] and [77] for the case of complex H-positive real matrices. Section 2.4 discusses the numerical range condition for H-positive real matrices, compare [77], where this was introduced for the H-dissipative case. Using this we discus the uniqueness and stability of invariant maximal semidefinite subspaces, the construction of which was aready discussed in [19]. Finally, in subsection 2.5.1 the real case is discussed, see also Theorem 1.1.5 and Example 1.1.6.

2.2 Jordan chains and nondegeneracy

We say that a chain of vectors $x_1, \ldots, x_k$ in $\mathbb{C}^n$ is a *Jordan chain* of an $n \times n$ matrix A corresponding to its eigenvalue λ if $x_1 \neq 0$ and

$$Ax_1 = \lambda x_1, \quad Ax_j = \lambda x_j + x_{j-1}, \quad j = 2, 3, \ldots, k.$$

The natural number k is called the *length* of the Jordan chain $x_1, \ldots, x_k$.

A subspace $\mathcal{M}$ is said to be *nondegenerate* with respect to the *indefinite inner product* $[.,.]$ if $x \in \mathcal{M}$ and $[x, y] = 0$ for all $y \in \mathcal{M}$ implies that $x = 0$. Otherwise, $\mathcal{M}$ is *degenerate*. Alternatively, one can show that $\mathcal{M}$ is nondegenerate if and only if $\mathcal{M}^{[\perp]}$ is a direct complement of $\mathcal{M}$ in $\mathbb{C}^n$. Here $\mathcal{M}^{[\perp]}$ denotes the orthogonal space of $\mathcal{M}$ with respect to the H-indefinite inner product. Further it is known (see e.g. [5] and [12]) that for any nondegenerate subspace $\mathcal{M}$ there exist subspaces $\mathcal{M}_-$ and $\mathcal{M}_+$ of $\mathcal{M}$ such that $\mathcal{M}_-$ is negative, $\mathcal{M}_+$ is positive and $\mathcal{M}_-[\dot{+}]\mathcal{M}_+ = \mathcal{M}$, where $[\dot{+}]$ means the orthogonal direct sum, with respect to the indefinite inner product. A constructive proof can be found in [77], Proposition 1.0.3.

Lemma 2.2.1 *Let A be an H-positive real matrix. Then $Re[Ax, x] = 0$ if and only if $Ax = -A^{[*]}x$, for $x \in \mathbb{C}^n$.*

Proof. Note that $[iAx, x] = i\langle HAx, x\rangle = i[Ax, x]$. Thus, $\mathrm{Re}[Ax, x] = \mathrm{Im}[iAx, x]$ and iA is H-dissipative. Therefore by Lemma 2.1.1 in [77], see also Theorem II.2.15 in [5], it follows that $\mathrm{Im}[iAx, x] = 0$ if and only if $iAx = (iA)^{[*]}x = -iA^{[*]}x$, that is, if and only if $Ax = -A^{[*]}x$. Thus, $\mathrm{Re}[Ax, x] = 0$ if and only if $Ax = -A^{[*]}x$. □

Lemma 2.2.2 *Let $x_1, \ldots, x_n$ be a Jordan chain of the H-positive real matrix A corresponding to its pure imaginary eigenvalue λ with $\lambda \in i\mathbb{R}$. Let $m = [\frac{n+1}{2}]$. Then $A^{[*]}x_j = -Ax_j$ for $j = 1, \ldots, m$.*

Proof. Let $x_0 = 0$ and let $k \in \{0, 1, \ldots, [\frac{n-1}{2}]\}$. Take note that $[\frac{n-1}{2}] + 1 = [\frac{n+1}{2}] = m$ for all $n = 1, 2, \ldots$. We prove by an induction argument that $\mathrm{Re}[Ax_j, x_j] = 0$ for all $j \in \{0, 1, \ldots, [\frac{n-1}{2}] + 1\}$. Note that for $j = 0$, we have $\mathrm{Re}[Ax_0, x_0] = \mathrm{Re}[0, 0] = 0$. Now assume that $\mathrm{Re}[Ax_j, x_j] = 0$ for all $j \in \{0, 1, \ldots, k\}$, where $k \in \{0, 1, \ldots, [\frac{n-1}{2}]\}$. Then by Lemma 2.2.1 we have $A^{[*]}x_j = -Ax_j$ for $j = 0, 1, \ldots, k$. Thus,

$$[x_{k+1}, x_k] = [(A - \lambda)x_{k+2}, x_k] = \ldots = [(A - \lambda)^k x_{2k+1}, x_k] = [x_{2k+1}, ((A - \lambda)^k)^{[*]}x_k].$$

Also,

$$\begin{aligned}
((A - \lambda)^k)^{[*]}x_k &= (A^{[*]} + \lambda)^k x_k \\
&= -(A^{[*]} + \lambda)^{k-1}(A - \lambda)x_k \\
&= -(A^{[*]} + \lambda)^{k-1}x_{k-1} \\
&= (-1)^2(A^{[*]} + \lambda)^{k-2}(A - \lambda)x_{k-1} \\
&= (-1)^2(A^{[*]} + \lambda)^{k-2}x_{k-2} \\
&= \vdots \\
&= (-1)^{k-1}(A^{[*]} + \lambda)x_1 \\
&= (-1)^k(A - \lambda)x_1 = 0.
\end{aligned}$$

Thus, $[x_{k+1}, x_k] = 0$. Therefore,

$$\begin{aligned}
\mathrm{Re}[Ax_{k+1}, x_{k+1}] &= \mathrm{Re}[(\lambda)x_{k+1} + x_k, x_{k+1}] \\
&= \mathrm{Re}(\lambda[x_{k+1}, x_{k+1}] + [x_k, x_{k+1}]) \\
&= \mathrm{Re}(\lambda[x_{k+1}, x_{k+1}]) = 0,
\end{aligned}$$

since $[x_{k+1}, x_{k+1}] \in \mathbb{R}$, because $[x_{k+1}, x_{k+1}] = \langle Hx_{k+1}, x_{k+1}\rangle$ and H is selfadjoint. This completes the induction argument. Hence from Lemma 2.2.1, $A^{[*]}x_j = -Ax_j$ for all $j \in \{0, 1, \ldots, [\frac{n-1}{2}] + 1\}$. □

Theorem 2.2.3 *Let $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ be two Jordan chains of an H-positive real matrix A corresponding to its pure imaginary eigenvalue $\lambda \in i\mathbb{R}$. Let $m_1 = [\frac{n_1+1}{2}]$ and $x_0 = y_0 = 0$. Then,*

$$[x_k, y_l] = -[x_{k-1}, y_{l+1}]$$

for all $k \in \{1, \ldots, m_1\}$ and $l \in \{0, 1, \ldots, n_2 - 1\}$.

Proof. We first justify the choice of $x_0 = y_0 = 0$ by noting that

$$[x_1, y_l] = [x_1, (A-\lambda)y_{l+1}] = [(A-\lambda)^{[*]}x_1, y_{l+1}] = -[(A-\lambda)x_1, y_{l+1}] = [0, y_{l+1}] = 0.$$

Therefore the choice of $x_0 = 0$ (and by interchanging the roles of x_k and y_l), $y_0 = 0$ makes sense.

It follows from Lemma 2.2.2 that $A^{[*]}x_j = -Ax_j$ for $j = 1, 2, \ldots, m_1$. Thus for $k \in \{1, 2 \ldots, m_1\}$ it holds that

$$[y_l, x_k] = [(A-\lambda)y_{l+1}, x_k] = [y_{l+1}, (A-\lambda)^{[*]}x_k] = -[y_{l+1}, (A-\lambda)x_k] = -[y_{l+1}, x_{k-1}]$$

By interchanging the roles of x_k and y_l it follows that $[x_k, y_l] = -[x_{k-1}, y_{l+1}]$. □

Let $n_1 = n_2 = n$ in Theorem 2.2.3. Then, since $[x_0, y_j] = [0, y_j] = 0$ for all $j = 0, 1, \ldots, n-1$, it follows that

$$[x_k, y_l] = -[x_{k-1}, y_{l+1}] = (-1)^2[x_{k-2}, y_{l+2}] = \ldots = (-1)^{k-1}[x_1, y_{l+k-1}] = (-1)^k[x_0, y_{l+k}] = 0,$$

for all $l + k \leq n$. If moreover n is even, then $m = \frac{n}{2}$ and $[x_k, y_l] = -[x_{k-1}, y_{l+1}]$ for $k = 1, 2, \ldots, \frac{n}{2}$ and $l = 0, 1, \ldots, n-1$. Furthermore it follows from Theorem 2.2.3 that the *Gram matrix* has the following form

$$\begin{bmatrix} [x_1,y_1] & \ldots & [x_n,y_1] \\ \vdots & \ddots & \vdots \\ [x_1,y_n] & \ldots & [x_n,y_n] \end{bmatrix} = \left[\begin{array}{ccc|ccc} & & & & & \bar{a} \\ & & & & & -\bar{a} \\ & 0 & & & \iddots & \\ & & & & (-1)^{\frac{n-2}{2}}\bar{a} & \iddots \\ \hline & & (-1)^{\frac{n-2}{2}}a & & & \\ & \iddots & & & * & \\ & -a & & & & \\ a & & \iddots & & & \end{array}\right], \tag{2.1}$$

where the horizontal and vertical lines divide the matrix in four submatrices of equal size.

When n is odd, the following Gram matrix structure appears, where the left upper block is of size $(n+1)/2$

$$\begin{bmatrix} [x_1,y_1] & \ldots & [x_n,y_1] \\ \vdots & \ddots & \vdots \\ [x_1,y_n] & \ldots & [x_n,y_n] \end{bmatrix} = \left[\begin{array}{ccc|ccc} & & & & & a \\ & & & & -a & \\ & 0 & & & \iddots & \\ & & & (-1)^{\frac{n-3}{2}}a & & \\ \hline & & (-1)^{\frac{n-1}{2}}a & & & \iddots \\ & \iddots & & & * & \\ & -a & & & & \\ a & & \iddots & & & \end{array}\right]. \tag{2.2}$$

In both cases there are zero entries in the upper left triangle. After dividing the matrix in four blocks as in the picture, the lower left block and upper right block have an alternating *Hankel structure*. That is, entries lying on the same skew diagonal and belonging to the same block have the same absolute value. In the odd case the main skew diagonal picks out one entry from the upper left block. The central entry, i.e. $[x_{\frac{n+1}{2}}, y_{\frac{n+1}{2}}]$, has the same modulus value as elsewhere on the main skew diagonal. In the lower right block no particular structure is identified. If $x_k = y_k$ for $k = 1, \ldots, n$, then in the odd case the central entry $[x_{\frac{n+1}{2}}, y_{\frac{n+1}{2}}]$ is a nonzero real number - and thus in this case the main skew diagonal consists of real numbers only.

When the lengths of the two chains in Theorem 2.2.3 are different (say $n_1 > n_2$), then the matrix of indefinite scalar products has the following form:

$$\begin{bmatrix} [x_1,y_1] & \ldots & [x_{n_1},y_1] \\ \vdots & \ddots & \vdots \\ [x_1,y_{n_2}] & \ldots & [x_{n_1},y_{n_2}] \end{bmatrix} = \left[\begin{array}{ccc|ccc} & & & & & b \\ & & & & -b & \\ & 0 & & & \iddots & \\ & & & & & \\ & & & (-1)^{\frac{n-2}{2}}b & \iddots & \\ \hline & & (-1)^{\frac{n-2}{2}}a & & & \\ & & & & & \\ & \iddots & & & * & \\ & -a & & & & \\ a & & \iddots & & & \end{array}\right]. \tag{2.3}$$

Corollary 2.2.4 *Let $x_1, \ldots, x_l$ be a Jordan chain of the H-positive real matrix A, corresponding to a pure imaginary eigenvalue λ of A. Then $l \leq 2\kappa + 1$, where κ is the number of negative eigenvalues of H.*

Proof. Consider the Jordan chain $\{x_1, x_2, \ldots, x_l\}$ $(l \leq n)$. Then this set is linearly independent. Now $[x_j, x_k] = 0$ for all $j + k \leq l$ by (2.1) and (2.2). Assume that $l > 2\kappa + 1$ and let $X = \mathrm{Sp}\{x_1, x_2, \ldots, x_{\frac{l-1}{2}}\}$, i.e $[x_j, x_k] = 0$ for x_j, x_k in the basis of X. Thus X is a *neutral subspace* with dimension $\frac{l-1}{2}$. Then we have $\dim X = \frac{l-1}{2} > \kappa$. According to Theorem 2.3.2 in [24] this is impossible, since κ is the maximal dimension of any nonpositive subspace (X being neutral it is nonpositive). We conclude $l \leq 2\kappa + 1$. □

Corollary 2.2.5 *Let $x_1, x_2, \ldots, x_n$ be a Jordan chain of the H-positive real matrix A, corresponding to the pure imaginary eigenvalue λ of A. Then the eigenvector x_1 is orthogonal to $Ker(A - \lambda)^{n-1}$.*

Proof. Let $y \in \mathrm{Ker}(A-\lambda)^{n-1}$. Thus,

$$\begin{aligned} 0 &= [x_n, (A-\lambda)^{n-1}y] \\ &= [(A-\lambda)^{[*]}x_n, (A-\lambda)^{n-2}y] \\ &= -[(A-\lambda)x_n, (A-\lambda)^{n-2}y] \\ &= -[x_{n-1}, (A-\lambda)^{n-2}y] \\ &= -[(A-\lambda)^{[*]}x_{n-1}, (A-\lambda)^{n-3}y] \\ &= (-1)^2[x_{n-2}, (A-\lambda)^{n-3}y] \\ &= \vdots \\ &= (-1)^{n-1}[x_1, (A-\lambda)^0 y] \\ &= (-1)^{n-1}[x_1, y]. \end{aligned}$$

Thus, $[x_1, y] = 0$ for all $y \in \mathrm{Ker}(A-\lambda)^{n-1}$, hence $x_1 \perp \mathrm{Ker}(A-\lambda)^{n-1}$. □

Corollary 2.2.6 *Let $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ be two Jordan chains of the H-positive real matrix A, both corresponding to the pure imaginary eigenvalue λ of A. Let $m_1 = [\frac{n_1+1}{2}]$ and let $m_2 = [\frac{n_2+1}{2}]$. If $n_1 \neq n_2$, then $Sp\{x_1, x_2, \ldots, x_{m_1}\}[\perp]Sp\{y_1, y_2, \ldots, y_{m_2}\}$.*

Proof. From (2.3) it follows that $[x_k, y_l] = 0$ for all $k = 1, \ldots, m_1$ and $l = 1, \ldots, m_2$. Thus, $\mathrm{Sp}\{x_1, x_2, \ldots, x_{m_1}\}[\perp]\mathrm{Sp}\{y_1, y_2, \ldots, y_{m_2}\}$. □

Theorem 2.2.7 *Let $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ be two Jordan chains of the H-positive real matrix A, corresponding to the eigenvalues λ and μ respectively, with $\lambda \neq \mu$. Let $m_1 = [\frac{n_1+1}{2}]$. If λ is purely imaginary, then $[x_k, y_l] = 0$ for all $k \in \{1, 2, \ldots, m_1\}$ and $l \in \{1, 2, \ldots, n_2\}$.*

Proof. From Lemma 2.2.2 it follows that $A^{[*]}x_k = -Ax_k$, $k \in \{1, 2, \ldots, m_1\}$. Let $x_0 = y_0 = 0$ and also $y_{-1} = 0$. Fix $j \in \{0, 1, \ldots, m_1 + n_2 - 1\}$. Assume that $[x_k, y_l] = 0$ for all pairs $k, l \in \mathbb{N} \cup \{0\}$ such that $k + l \leq j$, $k \leq m_1$ and $l \leq n_2$. Fix such a pair k, l. If $k \leq m_1 - 1$, then

$$\begin{aligned} \mu[y_l, x_{k+1}] &= [Ay_l - y_{l-1}, x_{k+1}] \\ &= [Ay_l, x_{k+1}] - [y_{l-1}, x_{k+1}] \\ &= [y_l, A^{[*]}x_{k+1}] \\ &= -[y_l, Ax_{k+1}] \\ &= -[y_l, \lambda x_{k+1} + x_k] \\ &= -\overline{\lambda}[y_l, x_{k+1}] - [y_l, x_k] \\ &= \lambda[y_l, x_{k+1}]. \end{aligned}$$

Since $\lambda \neq \mu$ it follows that $[y_l, x_{k+1}] = 0$. If $l \leq n_2 - 1$, then it follows similarly that $\mu[y_{l+1}, x_k] = \lambda[y_{l+1}, x_k]$. Thus, $[y_{l+1}, x_k] = 0$. Hence $[x_k, y_l] = 0$ for all pairs $k, l \in \mathbb{N} \cup \{0\}$, such that $k + l \leq j + 1$, $k \leq m_1$ and $l \leq n_2$. By induction on j it follows that $[x_k, y_l] = 0$ for all pairs $k, l \in \mathbb{N} \cup \{0\}$ such that $k \leq m_1$ and $l \leq n_2$. □

If μ is purely imaginary as well, then the roles of the first and second Jordan chains can be reversed. In this case the Gram matrix has the following structure

$$\begin{bmatrix} [x_1, y_1] & \dots & [x_{n_1}, y_1] \\ \vdots & \ddots & \vdots \\ [x_1, y_{n_2}] & \dots & [x_{n_1}, y_{n_2}] \end{bmatrix} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & * \end{array}\right]. \tag{2.4}$$

If μ is not purely imaginary, then

$$\begin{bmatrix} [x_1, y_1] & \dots & [x_{n_1}, y_1] \\ \vdots & \ddots & \vdots \\ [x_1, y_{n_2}] & \dots & [x_{n_1}, y_{n_2}] \end{bmatrix} = \left[\begin{array}{c|c} 0 & * \end{array}\right]. \tag{2.5}$$

Summarizing, the following corollary holds.

Corollary 2.2.8 *Let $x_1, x_2, \dots, x_{n_1}$ and $y_1, y_2, \dots, y_{n_2}$ be two Jordan chains of the H-positive real matrix A, corresponding to the eigenvalues λ and μ of A respectively, with $\lambda \neq \mu$. Let $m_1 = [\frac{n_1+1}{2}]$. If $\lambda \in i\mathbb{R}$, then $Sp\{x_1, x_2, \dots, x_{m_1}\}[\perp]Sp\{y_1, y_2, \dots, y_{n_2}\}$.*

We denote by $\mathcal{R}(A, \Lambda)$ the spectral subspace of A corresponding to the set $\Lambda \in \mathbb{C}$.

Lemma 2.2.9 *Let A be any matrix. For every subset Λ of $\mathbb{C}$ we have $\mathcal{R}(A, \Lambda) = \mathcal{R}(iA, i\Lambda)$.*

Proof.

$$\begin{aligned} \mathcal{R}(A, \Lambda) &= \mathrm{Sp}\{\mathrm{Ker}(A - \lambda)^k \quad | \quad \lambda \in \sigma(A) \cap \Lambda,\ k = 1, 2, \dots\} \\ &= \mathrm{Sp}\{\mathrm{Ker}(iA - i\lambda)^k \quad | \quad i\lambda \in \sigma(iA) \cap i\Lambda,\ k = 1, 2, \dots\} \\ &= \mathcal{R}(iA, i\Lambda) \end{aligned}$$

□

Let A be an H-positive real matrix and assume $\lambda \in \sigma(A)$. Then iA is H-dissipative and $i\lambda \in \sigma(iA)$. Thus, by Corollary 2.1.3 in [77], $\mathcal{R}(iA, \{i\lambda, \overline{i\lambda}\})$ is nondegenerate. Also by Lemma 2.2.9 with $\Lambda = \{\lambda, -\overline{\lambda}\}$ we get that $\mathcal{R}(A, \{\lambda, -\overline{\lambda}\}) = \mathcal{R}(iA, \{i\lambda, \overline{i\lambda}\})$. We therefore conclude that:

Corollary 2.2.10 *If A is H-positive real and $\lambda \in \mathbb{C}$ is any eigenvalue of A, then $\mathcal{R}(A, \{\lambda, -\overline{\lambda}\})$ is nondegenerate.*

Lemma 2.2.11 *Let A be H-positive real and let λ be an eigenvalue of A and $\{x_1, x_2, \dots, x_k\}$ a Jordan chain of A which corresponds to λ. Let $m = [\frac{k+1}{2}]$.*

*(i) If $\lambda \in \mathbb{C}_{left}$, then $Sp\{x_1, x_2, \dots, x_k\}$ is an A-*invariant nonpositive subspace.

*(ii) Let $\lambda \in i\mathbb{R}$ and k be even, then $Sp\{x_1, x_2, \dots, x_m\}$ is an A-*invariant neutral subspace.

(iii) Let $\lambda \in i\mathbb{R}$ and k be odd, then $Sp\{x_1, x_2, \dots, x_{m-1}\}$ is an A-invariant neutral subspace.

*(iv) Let $\lambda \in i\mathbb{R}$ and k be odd. If $[x_m, x_m] \geq 0$, then $Sp\{x_1, x_2, \dots, x_m\}$ is an A-*invariant nonnegative subspace.

(v) Let $\lambda \in i\mathbb{R}$ and let k be odd. If $[x_m, x_m] \leq 0$, then $Sp\{x_1, x_2, \ldots, x_m\}$ is an A-invariant nonpositive subspace.

(vi) $Ker(A-\lambda)^{2\kappa+2} = Ker(A-\lambda)^{2\kappa+1}$.

Proof.

(i) Since iA is H-dissipative and $i\lambda \in \mathbb{C}_{low}$ it follows from Lemma 2.3.2 (i) in [77] that $\mathrm{Sp}\{x_1, x_2, \ldots, x_k\}$ is an iA-invariant (also A-invariant) nonpositive subspace.

(ii) If $\lambda = i\mu$, with $\mu \in \mathbb{R}$, then $i\lambda = i(i\mu) = -\mu \in \mathbb{R}$, hence from Lemma 2.3.2 (ii) in [77], it follows that $\mathrm{Sp}\{x_1, x_2, \ldots, x_m\}$ is an A-invariant neutral subspace.

(iii) If $\lambda = i\mu$, with $\mu \in \mathbb{R}$, then $i\lambda = -\mu \in \mathbb{R}$, hence from Lemma 2.3.2 (iii) in [77], it follows that $\mathrm{Sp}\{x_1, x_2, \ldots, x_{m-1}\}$ is an A-invariant neutral subspace.

(iv) Let $\tilde{x}_j = (-i)^{j-1}x_j$ for $j = 1, 2, \ldots, m$. Note that

$$\begin{aligned}(iA - i\lambda)(\tilde{x}_j) &= (iA - i\lambda)((-i)^{j-1}x_j) \\ &= (i)(-i)^{j-1}(A-\lambda)x_j \\ &= (-i)^{j-2}x_{j-1} \\ &= \tilde{x}_{j-1} \quad \text{for} \quad j = 2, 3, \ldots, m,\end{aligned}$$

so that $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_m\}$ is a Jordan chain of iA corresponding to the eigenvalue $i\lambda$ of iA. Now, since $\lambda \in i\mathbb{R}$ and

$$[\tilde{x}_m, \tilde{x}_m] = [(-i)^{m-1}x_m, (-i)^{m-1}x_m] = [x_m, x_m] \geq 0,$$

it follows from Lemma 2.3.2 (iv) in [77] that $\mathrm{Sp}\{x_1, x_2, \ldots, x_m\} = \mathrm{Sp}\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_m\}$ is an A-invariant nonnegative subspace.

(v) Proof is as in (iv).

(vi) Let $x \in \mathrm{Ker}(A-\lambda)^{2\kappa+1}$, then $(A-\lambda)^{2\kappa+2}x = (A-\lambda)(A-\lambda)^{2\kappa+1}x = 0$. Thus $\mathrm{Ker}(A-\lambda)^{2\kappa+1} \subseteq \mathrm{Ker}(A-\lambda)^{2\kappa+2}$. Conversely, let $x \in \mathrm{Ker}(A-\lambda)^{2\kappa+2}$ and assume that $x \notin \mathrm{Ker}(A-\lambda)^{2\kappa+1}$. Then the set $\{(A-\lambda)^{2\kappa+1}x, (A-\lambda)^{2\kappa}x, \ldots, (A-\lambda)^{\kappa+1}x\}$ consists of the first $\kappa+1$ vectors in a Jordan chain of (even) length $2\kappa+2$. The subspace generated by this linear independent set is of dimension $\kappa+1$ and is neutral by equation (2.1), contradicting the fact that such a subspace of dimension larger than κ does not exist. Hence $\mathrm{Ker}(A-\lambda)^{2\kappa+2} = \mathrm{Ker}(A-\lambda)^{2\kappa+1}$. □

Proposition 2.2.12 *Let A be H-positive real, $\lambda \in i\mathbb{R}$ an eigenvalue of A. Let $\{J_1, J_2, \ldots, J_{2\kappa+1}\}$ be a* Jordan basis *for $\mathcal{R}(A, \{\lambda\})$ where J_j is the subset of all Jordan chains of length j. Then $\mathcal{R}(A, \{\lambda\})$ is nondegenerate if and only if SpJ_j is nondegenerate for all $j = 1, 2, \ldots, 2\kappa+1$.*

Proof. We know iA is H-dissipative and $i\lambda \in \mathbb{R}$ is an eigenvalue of iA. Let $x_1, x_2, \ldots, x_j$ be a Jordan chain in J_j. Then $\{x_1, -ix_2, \ldots, (-i)^{j-1}x_j\} \in J_j'$ where J_j' is the set of all Jordan chains of iA corresponding to $i\lambda$ of length j. Thus, $\mathrm{Sp}J_j = \mathrm{Sp}J_j'$ and also

$\mathcal{R}(A,\{\lambda\}) = \mathcal{R}(iA,\{i\lambda\})$. Therefore the result follows directly from Proposition 2.3.3 in [77]. □

Let $j \in \{1,3,\ldots,2\kappa+1\}$ and let $x_k \in J_j \backslash \mathrm{Ker}(A-\lambda)^{j-1}$, $k = 1,2,\ldots$. If $z_1, z_2, \ldots, z_j$ is a Jordan chain of matrix A corresponding to its eigenvalue λ, then

$$z_1 = (A-\lambda)z_2 = (A-\lambda)^2 z_3 = \ldots = (A-\lambda)^{j-1} z_j$$

and

$$(A-\lambda)^{j-1} z_k = (A-\lambda)^{j-2} z_{k-1} = \ldots = (A-\lambda)^{j-k} z_1 = (A-\lambda)^{j-k-1}(A-\lambda)z_1 = 0$$

for all $1 \le k \le j-1$. Thus in each chain $z_1, z_2, \ldots, z_j$ only the last element z_j belongs to $J_j \backslash \mathrm{Ker}(A-\lambda)^{j-1}$. There are only a finite number, say r_j, of such chains in J_j, thus only a finite number of choices for $z_j \in J_j \backslash \mathrm{Ker}(A-\lambda)^{j-1}$. Thus the set $\{x_1, x_2, \ldots, x_{r_j}\}$ is the set of all the last elements of the Jordan chains of J_j.

Let $m = \frac{j+1}{2}$ and put

$$y_1 = (A-\lambda)^{m-1} x_1, \quad y_2 = (A-\lambda)^{m-1} x_2, \ \ldots, \ y_{r_j} = (A-\lambda)^{m-1} x_{r_j}. \tag{2.6}$$

Since x_1 is the j^{th} element of a Jordan chain $z_1, z_2, \ldots, z_j$, we have

$$(A-\lambda)x_1 = z_{j-1}, (A-\lambda)^2 x_1 = (A-\lambda)z_{j-1} = z_{j-2}, \ldots, (A-\lambda)^{m-1} x_1 = z_{j-m+1}.$$

Thus, we conclude that $y_1 = (A-\lambda)^{m-1} x_1 = z_{j-m+1} = z_m$, which is a vector in the middle of the Jordan chain $z_1, z_2, \ldots, z_j$.

Lemma 2.2.13 *Let A be H-positive real, and $\lambda \in i\mathbb{R}$ an eigenvalue of A. For $j \in \{1,3,\ldots,2\kappa+1\}$ define $\{y_1, y_2, \ldots, y_{r_j}\}$ as in (2.6). Then SpJ_j is nondegenerate if and only if $Sp\{y_1, y_2, \ldots, y_{r_j}\}$ is nondegenerate.*

Proof. We know iA is H-dissipative and $i\lambda \in \mathbb{R}$ is an eigenvalue of iA. Let J_j' be as in the proof of Proposition 2.2.12. If $\{z_1, z_2, \ldots, z_j\}$ is a Jordan chain in J_j then $\{z_1, -iz_2, \ldots, (-i)^{j-1} z_j\}$ is a Jordan chain in J_j'. Let $x_k' = (-i)^{j-1} x_k$ for $k = 1,2,\ldots$. Now consider,

$$\begin{aligned}
y_1' &= (iA - i\lambda)^{m-1} x_1', \quad \text{where} \quad m = \frac{j+1}{2} \\
&= (i)^{m-1}(A-\lambda)^{m-1}(-i)^{j-1} x_1 \\
&= (-1)^{j-1}(i)^{m+j-2}(A-\lambda)^{m-1} x_1 \\
&= (i)^{\frac{3(j-1)}{2}} y_1 \\
y_2' &= (i)^{\frac{3(j-1)}{2}} y_2 \\
&\vdots
\end{aligned}$$

Since $\mathrm{Sp}\{y_1', y_2', \ldots, y_{r_j}'\} = \mathrm{Sp}\{y_1, y_2, \ldots, y_{r_j}\}$, our result follows from Lemma 2.3.4 in [77]. □

Note that the Gram matrices associated with $y_{j,k}'$ (an element of the Jordan chain with respect to the eigenvalue $i\lambda$ of the H-dissipative matrix iA) and $y_{j,k}$ (an element of

the Jordan chain with respect to the pure imaginary eigenvalue λ of the H-positive real matrix A) are equal. That is $[y'_{j,k}, y'_{j,l}] = [y_{j,k}, y_{j,l}]$, since

$$\begin{aligned}[y'_{j,k}, y'_{j,l}] &= [i^{\frac{3(j-1)}{2}} y_{j,k}, i^{\frac{3(j-1)}{2}} y_{j,l}] \\ &= (i^{\frac{3(j-1)}{2}})(\overline{i^{\frac{3(j-1)}{2}}})[y_{j,k}, y_{j,l}] \\ &= [y_{j,k}, y_{j,l}]\end{aligned}$$

2.3 The construction of maximal invariant subspaces

Let A be an H-positive real matrix with corresponding eigenvalue $\lambda \in i\mathbb{R}$. Let $J(\lambda) = \{J(\lambda,1), \ldots, J(\lambda, 2\kappa+1)\}$ be a Jordan basis of $\mathcal{R}(A, \{\lambda\})$, where $J(\lambda, j)$ is the set of all elements of the Jordan basis that belong to some Jordan chain of length j. Let $n_{\lambda,j}$ be the number of Jordan chains in the set $J(\lambda, j)$ and let $x_{j,1}, x_{j,2}, \ldots, x_{j,n_{\lambda,j}}$ be the vectors in $J(\lambda, j) \backslash \mathrm{Ker}(A-\lambda)^{j-1}$. Thus each $x_{j,k}$ is the last element in one of the $n_{\lambda,j}$ chains in $J(\lambda, j)$. Then

$$J(\lambda, j) = \{(A-\lambda)^k x_{j,l} \mid k = 0, 1, \ldots, j-1 \text{ and } l = 1, 2, \ldots, n_{\lambda,j}\}.$$

Put $m_j = [\frac{j+1}{2}]$. Let $y_{j,k} = (A-\lambda)^{m_j-1} x_{j,k}$ for $k = 1, 2, \ldots, n_{\lambda,j}$. If $\mathrm{Sp}\{y_{j,1}, \ldots, y_{j,n_{\lambda,j}}\}$ is nondegenerate, then from Proposition 1.0.3 in [77] it follows that $\mathrm{Sp}\{y_{j,1}, \ldots, y_{j,n_{\lambda,j}}\} = \mathcal{M}_-(\lambda, j)[\dot{+}]\mathcal{M}_+(\lambda, j)$ where $\mathcal{M}_-(\lambda, j)$ and $\mathcal{M}_+(\lambda, j)$ are nonpositive and nonnegative subspaces, respectively. It also follows from the construction that an element, say u, of $\mathcal{M}_-(\lambda, j)$ (similarly for an element of $\mathcal{M}_+(\lambda, j)$) can be written as $u = \sum_{s=1}^{n_{\lambda,j}} g_s y_{j,s}$, for some choice of g_s.

We use the subspaces $\mathcal{M}_-(\lambda, j)$ and $\mathcal{M}_+(\lambda, j)$ to construct invariant maximal nonnegative and nonpositive subspaces as follows:

(i) First observe that

$$[\mathrm{Ker}(A-\lambda)^{m_j-1} \cap \mathrm{Sp}J(\lambda, j)] \cap \mathcal{M}_-(\lambda, j) = \{0\},$$

because: if $u \in (\mathrm{Ker}(A-\lambda)^{m_j-1} \cap \mathrm{Sp}J(\lambda, j)) \cap \mathcal{M}_-(\lambda, j)$, then $u = \sum_{s=1}^{n_{\lambda,j}} g_s y_{j,s}$. Therefore $(A-\lambda)^{m_j-1} u = 0$ and thus $\sum_{s=1}^{n_{\lambda,j}} g_s (A-\lambda)^{m_j-1} y_{j,s} = 0$. Since $\{(A-\lambda)^{m_j-1} y_{j,s} \mid s = 1, \ldots, n_{\lambda,j}\}$ is linearly independent, it follows that $g_s = 0$ for $s = 1, 2, \ldots, n_{\lambda,j}$.

(ii) Let $x \in \mathrm{Ker}(A-\lambda)^{m_j-1} \cap \mathrm{Sp}J(\lambda, j)$. Then $x \in \mathrm{Sp}J(\lambda, j)$ means it can be written as a linear combination $x = \sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l} (A-\lambda)^k x_{j,l}$ and also since $x \in \mathrm{Ker}(A-\lambda)^{m_j-1}$ it follows that

$$\begin{aligned}\sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l} (A-\lambda)^k (A-\lambda)^{m_j-1} x_{j,l} &= (A-\lambda)^{m_j-1} x = 0, \text{ i.e.} \\ \sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l} (A-\lambda)^{k+m_j-1} x_{j,l} &= 0.\end{aligned}$$

However, $(A-\lambda)^{k+m_j-1}x_{j,l}=0$ for $k+m_j-1\geq j$, that is, for $k\geq j-m_j+1$. We may therefore write $0=\sum_{k=0}^{j-m_j}\sum_{l=1}^{n_{\lambda,j}}\beta_{k,l}(A-\lambda)^{k+m_j-1}x_{j,l}$. It then follows from the linear independence of the vectors in the Jordan basis that $\beta_{k,l}=0$ for all $k=0,1,\ldots,j-m_j$ and $l=1,2,\ldots,n_{\lambda,j}$. Thus,

$$x=\sum_{k=j-m_j+1}^{j-1}\sum_{l=1}^{n_{\lambda,j}}\beta_{k,l}(A-\lambda)^k x_{j,l}.$$

On the other hand, let $y\in\mathcal{M}_-(\lambda,j)$, then as before $y=\sum_{s=1}^{n_{\lambda,j}}g_s y_{j,s}$. Thus we have,

$$\begin{aligned}[x,y] &= [\sum_{k=j-m_j+1}^{j-1}\sum_{l=1}^{n_{\lambda,j}}\beta_{k,l}(A-\lambda)^k x_{j,l},\sum_{s=1}^{n_{\lambda,j}}g_s y_{j,s}]\\ &= \sum_{k=j-m_j+1}^{j-1}\sum_{l=1}^{n_{\lambda,j}}\sum_{s=1}^{n_{\lambda,j}}\beta_{k,l}\overline{g}_s[(A-\lambda)^k x_{j,l},y_{j,s}].\end{aligned}$$

Here $y_{j,s}=(A-\lambda)^{m_j-1}x_{j,s}$ and $x_{j,s}$ is the last element of the s^{th} Jordan chain i.e. $y_{j,s}$ is the middle term of the s^{th} chain. Since $x_{j,l}$ is the last element of the l^{th} chain, $(A-\lambda)^k x_{j,l}$ precedes it in the same chain. It follows from Corollary 2.2.6 that $[(A-\lambda)^k x_{j,l},y_{j,s}]=0$. Thus $[x,y]=0$.

We may therefore form the orthogonal direct sum:

$$\mathcal{N}_-(\lambda,j)=(\mathrm{Ker}(A-\lambda)^{m_j-1}\cap \mathrm{Sp}J(\lambda,j))[\dot{+}]\mathcal{M}_-(\lambda,j). \tag{2.7}$$

Similarly, using the same arguments, we may form the following direct sum:

$$\mathcal{N}_+(\lambda,j)=(\mathrm{Ker}(A-\lambda)^{m_j-1}\cap \mathrm{Sp}J(\lambda,j))[\dot{+}]\mathcal{M}_+(\lambda,j). \tag{2.8}$$

It remains to show that $\mathcal{N}_-(\lambda,j)$ and $\mathcal{N}_+(\lambda,j)$ are nonpositive and nonnegative subspaces respectively. We only prove the nonpositivity of $\mathcal{N}_-(\lambda,j)$; the nonnegativity of $\mathcal{N}_+(\lambda,j)$ follows similarly.

Let $x=u+v$ with $u\in\mathcal{M}_-(\lambda,j)$ and $v\in(\mathrm{Ker}(A-\lambda)^{m_j-1}\cap \mathrm{Sp}J(\lambda,j))$. Thus, $u=\sum_{s=1}^{n_{\lambda,j}}g_s y_{j,s}$ and $v=\sum_{k=0}^{j-1}\sum_{l=1}^{n_{\lambda,j}}\beta_{k,l}(A-\lambda)^k x_{j,l}$. Now, $[x,x]=[u+v,u+v]=[u,u]+[u,v]+[v,u]+[v,v]$. Clearly $[u,u]\leq 0$, since $u\in\mathcal{M}_-(\lambda,j)$ and $[u,v]=[v,u]=0$ from the orthogonal direct sum (see (ii)). We now prove that $[v,v]=0$. Let $v\in(\mathrm{Ker}(A-\lambda)^{m_j-1}\cap \mathrm{Sp}J(\lambda,j))$. As before we obtain

$$v=\sum_{k=j-m_j+1}^{j-1}\sum_{l=1}^{n_{\lambda,j}}\beta_{k,l}(A-\lambda)^k x_{j,l}.$$

The vector v is the linear combination of the first m_j-1 elements of the Jordan chains of length j in the Jordan basis. Now from Lemma 2.2.11 (iii) and Corollary 2.2.6, it follows that $[v,v]=0$. Hence $\mathcal{N}_-(\lambda,j)$ is a nonpositive subspace.

We prove for $x \in \mathcal{N}_-(\lambda, j) \cup \mathcal{N}_+(\lambda, j)$ that $A^{[*]}x = -Ax$. Let $x \in \mathcal{N}_-(\lambda, j)$. As before, $x = u + v$ with $u \in \mathcal{M}_-(\lambda, j)$ and $v \in (\mathrm{Ker}(A-\lambda)^{m_j-1} \cap \mathrm{Sp}J(\lambda, j))$. Now, $A^{[*]}x = A^{[*]}(u+v) = A^{[*]}u + A^{[*]}v$ and from Lemma 2.2.2 we have that,

$$A^{[*]}u = \sum_{s=1}^{n_{\lambda,j}} g_s A^{[*]} y_{j,s} = -\sum_{s=1}^{n_{\lambda,j}} g_s A y_{j,s} = -Au$$

and

$$\begin{aligned} A^{[*]}v &= \sum_{k=0}^{j-1}\sum_{l=1}^{n_{\lambda,j}} \beta_{k,l} A^{[*]}(A-\lambda)^k x_{j,l} \\ &= \sum_{k=0}^{j-1}\sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(-A(A-\lambda)^k x_{j,l}) \\ &= -Av. \end{aligned}$$

Thus, $A^{[*]}x = -Au - Av = -A(u+v) = -Ax$. Similarly it can be shown that $A^{[*]}x = -Ax$ for all $x \in \mathcal{N}_+(\lambda, j)$. Thus it holds for all $x \in \mathcal{N}_-(\lambda, j) \cup \mathcal{N}_+(\lambda, j)$.

Theorem 2.3.1 *Let A be H-positive real and $\lambda \in i\mathbb{R}$ an eigenvalue of A. Assume $\mathcal{R}(A, \{\lambda\})$ is nondegenerate. For any $j \in \{1, 3, \ldots, 2\kappa+1\}$ let $\mathcal{N}_-(\lambda, j)$ and $\mathcal{N}_+(\lambda, j)$ be as in (2.7) and (2.8). Then the subspace $\mathcal{N}_-(\lambda, j)$ is A-invariant and nonpositive, the subspace $\mathcal{N}_+(\lambda, j)$ is A-invariant and nonnegative and* $\dim \mathcal{N}_-(\lambda, j) + \dim \mathcal{N}_+(\lambda, j) = \dim SpJ(\lambda, j)$.

Proof. For λ a pure imaginary eigenvalue of the H-positive real matrix A, we know $i\lambda$ is a real eigenvalue of the H-dissipative matrix iA. We can therefore claim that $\mathcal{N}_-(\lambda, j)$ (and also $\mathcal{N}_+(\lambda, j)$) of the matrix A, equals $\mathcal{N}_-(i\lambda, j)$ (and $\mathcal{N}_+(i\lambda, j)$) of the matrix iA, respectively. This is true because one finds that for the matrix A (with eigenvalue $\lambda \in i\mathbb{R}$) and iA (with eigenvalue $i\lambda \in \mathbb{R}$) that

(i) $\mathrm{Sp}J(i\lambda, j) = \mathrm{Sp}J(\lambda, j)$

(ii) $\mathrm{Ker}(iA - i\lambda)^{m_j-1} = \mathrm{Ker}((i)^{m_j-1}(A-\lambda)^{m_j-1}) = \mathrm{Ker}(A-\lambda)^{m_j-1}$

(iii) $\mathcal{M}_-(\lambda, j) = \mathcal{M}_-(i\lambda, j)$, since

$$\begin{aligned} u \in \mathcal{M}_-(i\lambda, j) \Leftrightarrow u &= \sum_{s=1}^{n_{\lambda,j}} q_s y'_{j,s} \\ &= \sum_{s=1}^{n_{\lambda,j}} (i)^{\frac{3(j-1)}{2}} q_s y_{j,s}, \end{aligned}$$

where $y_{j,s}$ is as before (see (i)) and $y'_{j,s}$ is as in the proof of Lemma 2.2.13.

Therefore from Theorem 2.3.5 in [77] it follows that

$$\dim \mathcal{N}_-(\lambda, j) + \dim \mathcal{N}_+(\lambda, j) = \dim \mathrm{Sp}J(\lambda, j).$$

□

Furthermore,

(i) $\dim \mathrm{Sp}J(\lambda, j) = jn_{\lambda,j}$;

(ii) $\dim \mathcal{N}_-(\lambda, j) = (m_j - 1)n_{\lambda,j} + \dim \mathcal{M}_-(\lambda, j)$;

(iii) $\dim \mathcal{N}_+(\lambda, j) = (m_j - 1)n_{\lambda,j} + \dim \mathcal{M}_+(\lambda, j)$;

(iv) $\dim \mathcal{M}_-(\lambda, j) + \dim \mathcal{M}_+(\lambda, j) = n_{\lambda,j}$.

The maximality of the subspaces follows from the above properties.

For the even case $j = 2, 4, \ldots, 2\kappa$, let

$$\mathcal{N}(\lambda, j) = \mathcal{N}_-(\lambda, j) = \mathcal{N}_+(\lambda, j) = \mathrm{Ker}(A - \lambda)^{m_j} \cap \mathrm{Sp}J(\lambda, j). \tag{2.9}$$

From the proof of Theorem 2.3.1 we observe that for an H-positive real matrix A with eigenvalue $\lambda \in i\mathbb{R}$ (and iA being H-dissipative with eigenvalue $i\lambda \in \mathbb{R}$) that

$$\begin{aligned} \mathrm{Ker}(A - \lambda)^{m_j} &= \mathrm{Ker}(iA - i\lambda)^{m_j}, \\ \mathrm{Sp}J(\lambda, j) &= \mathrm{Sp}J(i\lambda, j). \end{aligned}$$

Therefore we can conclude that $\mathcal{N}(\lambda, j) = \mathcal{N}(i\lambda, j)$. Then according to Theorem 2.3.6 in [77], it follows that $\mathcal{N}(\lambda, j)$ is a neutral subspace. We use Lemma 2.2.11 as follows: if $v \in \mathcal{N}(\lambda, j) = \mathrm{Ker}(A - \lambda)^{m_j} \cap \mathrm{Sp}J(\lambda, j)$, then since $v \in \mathrm{Sp}J(\lambda, j)$ it follows that

$$v = \sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(A - \lambda)^k x_{j,l}$$

and because $v \in \mathrm{Ker}(A - \lambda)^{m_j}$, it follows that

$$\begin{aligned} \sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(A - \lambda)^k (A - \lambda)^{m_j} x_{j,l} &= 0, \text{ i.e.} \\ \sum_{k=0}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(A - \lambda)^{k+m_j} x_{j,l} &= 0. \end{aligned}$$

However, $(A-\lambda)^{k+m_j} x_{j,l} = 0$ for $k+m_j \geq j$, that is, for $k \geq j-m_j$. We may therefore write $0 = \sum_{k=0}^{j-m_j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(A - \lambda)^{k+m_j} x_{j,l}$. It therefore follows from the linear independence of the vectors in the Jordan basis that $\beta_{k,l} = 0$ for all $k = 0, 1, \ldots, j - m_j - 1$ and $l = 1, 2, \ldots, n_{\lambda,j}$. Thus,

$$v = \sum_{k=j-m_j}^{j-1} \sum_{l=1}^{n_{\lambda,j}} \beta_{k,l}(A - \lambda)^k x_{j,l}.$$

The vector v is the linear combination of the first m_j elements of the Jordan chains of length j in the Jordan basis. Now, from Lemma 2.2.11 (iii), it follows that $[v, v] = 0$. Hence $\mathcal{N}(\lambda, j)$ is a neutral subspace.

In particular it also follows from Lemma 2.3.6 in [77] that $\mathcal{N}_-(\lambda, j)$ is an A-invariant (also iA-invariant) nonpositive subspace and $\mathcal{N}_+(\lambda, j)$ is an A-invariant nonnegative subspace. From (i) and (ii) in the proof of Theorem 2.3.1 it follows that

$$\dim \mathcal{N}_-(\lambda, j) + \dim \mathcal{N}_+(\lambda, j) = \dim \mathrm{Sp}J(\lambda, j).$$

Also, as before $A^*x = -Ax$ for all $x \in \mathcal{N}(\lambda, j)$. The maximality of the subspaces also follows from similar arguments as before.

For odd and even j, it follows from the proof of Theorem 2.3.1 that

$$\begin{aligned} \mathcal{N}_-(\lambda, j) &= \mathcal{N}_-(i\lambda, j), \\ \mathcal{N}_+(\lambda, j) &= \mathcal{N}_+(i\lambda, j). \end{aligned}$$

With $\mathcal{N}_-(\lambda, j)$ and $\mathcal{N}_+(\lambda, j)$ as in (2.7) to (2.9), let

$$\mathcal{N}_-(\lambda) = \mathcal{N}_-(\lambda, 1)[\dot{+}]\mathcal{N}_-(\lambda, 2)[\dot{+}] \dots [\dot{+}]\mathcal{N}_-(\lambda, 2\kappa + 1), \tag{2.10}$$

$$\mathcal{N}_+(\lambda) = \mathcal{N}_+(\lambda, 1)[\dot{+}]\mathcal{N}_+(\lambda, 2)[\dot{+}] \dots [\dot{+}]\mathcal{N}_+(\lambda, 2\kappa + 1). \tag{2.11}$$

Thus we get that

$$\mathcal{N}_-(\lambda) = \mathcal{N}_-(i\lambda), \qquad \mathcal{N}_+(\lambda) = \mathcal{N}_+(i\lambda).$$

According to Corollary 2.2.6 the direct sums are indeed orthogonal direct sums. The subspaces $\mathcal{N}_-(\lambda)$ and $\mathcal{N}_+(\lambda)$ are nonpositive, respectively, nonnegative subspaces of maximal dimension within $\mathcal{R}(A, \{\lambda\})$. From the preceding discussion and Theorem 2.3.7 in [77] the following theorem holds:

Theorem 2.3.2 *Let $\lambda \in i\mathbb{R}$ be an eigenvalue of an H-positive real matrix A. Assume $\mathcal{R}(A, \{\lambda\})$ is nondegenerate. The subspace $\mathcal{N}_-(\lambda)$ from (2.10) is A-invariant nonpositive, the subspace $\mathcal{N}_+(\lambda)$ from (2.11) is A-invariant nonnegative and*

$$\dim \mathcal{N}_-(\lambda) + \dim \mathcal{N}_+(\lambda) = \dim \mathcal{R}(A, \{\lambda\}).$$

Moreover, $A^*x = -Ax$ for all $x \in \mathcal{N}_-(\lambda) \cup \mathcal{N}_+(\lambda)$.

For $\lambda_1, \lambda_2, \dots, \lambda_k$ purely imaginary eigenvalues of an H-positive real matrix A, let

$$\mathcal{N}_- = \mathcal{N}_-(\lambda_1)[\dot{+}]\mathcal{N}_-(\lambda_2)[\dot{+}] \dots [\dot{+}]\mathcal{N}_-(\lambda_k) \tag{2.12}$$

and

$$\mathcal{N}_+ = \mathcal{N}_+(\lambda_1)[\dot{+}]\mathcal{N}_+(\lambda_2)[\dot{+}] \dots [\dot{+}]\mathcal{N}_+(\lambda_k). \tag{2.13}$$

The subspaces $\mathcal{N}_-(\lambda_i)$ and $\mathcal{N}_+(\lambda_i)$ for $i = 1, 2, \dots, k$ are constructed as in (2.10) and (2.11). It follows from Corollary 2.2.8 that the direct sums are orthogonal direct sums. Hence the subspaces $\mathcal{N}_-$ and $\mathcal{N}_+$ are the same as the corresponding subspaces $\mathcal{N}_-$ and $\mathcal{N}_+$ as in [77], p.36.

Now let,

$$\mathcal{M}_- = \mathcal{N}_-[\dot{+}]\mathcal{R}(A, \mathbb{C}_{left}), \tag{2.14}$$

$$\mathcal{M}_+ = \mathcal{N}_+[\dot{+}]\mathcal{R}(A, \mathbb{C}_{right}). \tag{2.15}$$

By Lemma 2.2.9 it means that

$$\begin{aligned} \mathcal{M}_- &= \mathcal{N}_-[\dot{+}]\mathcal{R}(iA, \mathbb{C}_{low}), \\ \mathcal{M}_+ &= \mathcal{N}_+[\dot{+}]\mathcal{R}(iA, \mathbb{C}_{upp}). \end{aligned}$$

Thus, we get the same subspaces $\mathcal{M}_-$ and $\mathcal{M}_+$ as in [77] for real eigenvalues. Therefore it follows that $\mathcal{M}_-$ from (2.14) and $\mathcal{M}_+$ from (2.15) are A-invariant maximal nonpositive and nonnegative subspaces, respectively.

2.4 The numerical range condition

Let A be an H-positive real matrix and suppose $\lambda \in i\mathbb{R}$ is an eigenvalue of A. Then iA is H-dissipative and $i\lambda \in \mathbb{R}$. As before, let κ denote the number of negative eigenvalues of H. By Corollary 2.2.4, the maximum length of a Jordan chain of A corresponding to λ does not exceed $2\kappa + 1$. Consider a Jordan basis for $\mathcal{R}(A, \lambda)$ and split it into the sets $J(\lambda, j)$, where $J(\lambda, j)$ consists of the basis vectors belonging to Jordan chains of length j. Denote by $n_{\lambda,j}$ the number of chains of length j. Now denote the basis vectors in $J(\lambda, j)$ that are not in the set $\mathrm{Ker}(A - \lambda)^{j-1}$ by $\{x_{j;1}, \ldots, x_{j;n_{\lambda,j}}\}$. These are the vectors at the end of a Jordan chain of length j. Let m_j be $\frac{j+1}{2}$ in case j is odd, and $\frac{j}{2}$ in case j is even. Define as before

$$y_{j;k} = (A - \lambda)^{m_j - 1} x_{j;k},$$

which are the vectors in the middle of the chain. We first deal with the case of Jordan chains of odd length. Following the arguments of [77] p.50 in this case for (A, λ) as above, we let

$$CM_j = \begin{bmatrix} \langle Hy_{j;1}, y_{j;1}\rangle & \ldots & \langle Hy_{j;n_{\lambda,j}}, y_{j;1}\rangle \\ \vdots & & \vdots \\ \langle Hy_{j;1}, y_{j;n_{\lambda,j}}\rangle & \ldots & \langle Hy_{j;n_{\lambda,j}}, y_{j;n_{\lambda,j}}\rangle \end{bmatrix},$$

for $j = 1, 3, \ldots, 2\kappa + 1$. (Here CM stands for the *characteristic matrix*). Define

$$CM_{odd}(A, \lambda) = \mathrm{diag}\{CM_1, CM_3, \ldots, CM_{2\kappa+1}\}.$$

Let $n_{odd} = n_{\lambda,1} + n_{\lambda,3} + \ldots + n_{\lambda,2\kappa+1}$ and put

$$NR_{odd}(A, \lambda) = \{\langle CM_{odd}(A, \lambda)x, x\rangle | x \in \mathbb{C}^{n_{odd}}, x \neq 0\}.$$

(Here NR stands for the *numerical range*).

By defining $x'_{j,k}$ and $y'_{j,k}$ as in the proof of Lemma 2.2.13 and using that $[y'_{j,k}, y'_{j,l}] = [y_{j,k}, y_{j,l}]$ we have that our $CM_{odd}(A, \lambda)$ is the same matrix as the matrix $CM_{odd}(iA, i\lambda)$ in [77], where iA is H-dissipative and $i\lambda \in \mathbb{R}$. Similarly, $NR_{odd}(A, \lambda)$ corresponds to $NR_{odd}(iA, i\lambda)$ in [77]. Recall from [77], Section 3.1.1, that $CM_{odd}(iA, i\lambda)$ is Hermitian and invertible and $NR_{odd}(iA, i\lambda)$ is independent of the choice of the Jordan basis. Hence the same properties hold for $CM_{odd}(A, \lambda)$ and $NR_{odd}(A, \lambda)$. We may therefore define the *odd numerical range condition* for (A, H) as follows:

Definition 2.4.1 *Let A be H-positive real, then the pair (A, H) is said to satisfy the odd numerical range condition if $0 \notin NR_{odd}(A, \lambda)$ for all $\lambda \in i\mathbb{R} \cap \sigma(A)$. (Thus, (A, H) has the odd numerical range condition if and only if (iA, H) has the odd numerical range condition for the dissipative matrices in* [77]*).*

For even length chains, let $j, k \in \{2, 4, \ldots, 2\kappa\}$ and put

$$CM_{j,k} = \begin{bmatrix} \langle H(A-\lambda)y_{j;1}, y_{k;1}\rangle & \ldots & \langle H(A-\lambda)y_{j;n_{\lambda,j}}, y_{k;1}\rangle \\ \vdots & & \vdots \\ \langle H(A-\lambda)y_{j;1}, y_{k;n_{\lambda,k}}\rangle & \ldots & \langle H(A-\lambda)y_{j;n_{\lambda,j}}, y_{k;n_{\lambda,k}}\rangle \end{bmatrix},$$

where the vectors $y_{j,1}, y_{j,2}, \ldots, y_{j,n_{\lambda,j}}$ and the vectors $y_{k,1}, y_{k,2}, \ldots, y_{k,n_{\lambda,k}}$ are defined as before. Let

$$CM_{even}(A, \lambda) = \begin{bmatrix} CM_{2,2} & CM_{4,2} & \ldots & CM_{2\kappa,2} \\ \vdots & & & \vdots \\ CM_{2,2\kappa} & CM_{4,2\kappa} & \ldots & CM_{2\kappa,2\kappa} \end{bmatrix}.$$

With $x'_{j,k}$ as in the proof of Lemma 2.2.13, we now find with j, k even and $m = \frac{j}{2}$ that

$$\begin{aligned} y'_{j,l} &= (-1)i^{\frac{3j-4}{2}} y_{j,l} \\ &= i^2 i^{\frac{3j-4}{2}} y_{j,l} \\ &= i^{\frac{3j}{2}} y_{j,l}, \end{aligned}$$

and

$$y'_{k,l} = i^{\frac{3k}{2}} y_{k,l}.$$

It then follows that

$$[(iA - i\lambda)y'_{j,l}, y'_{k,m}] = i^{\frac{9k+3j}{2}} [y_{j,l-1}, y_{k,m}]$$

with respect to the two new Jordan chains $y'_{j,1}, y'_{j,2}, \ldots, y'_{j,n_{\lambda,j}}$ and $y'_{k,1}, y'_{k,2}, \ldots, y'_{k,n_{\lambda,k}}$. The characteristic matrix for $(iA, i\lambda)$ then becomes

$$CM'_{j,k} = i^{\frac{9k+3j}{2}} CM_{j,k}.$$

In particular for $j = k$ we see that

$$CM'_{j,j} = CM_{j,j}.$$

From equation (2.5) in [77] it follows that $CM'_{j,k} = 0$ for $j < k$, hence it also follows that $CM_{j,k} = 0$ for $j < k$. Thus $CM_{even}(A, \lambda)$ is a block upper triangular matrix. Using that $CM'_{j,j}$ is invertible (as is proved in Proposition 3.1.4 in [77]), it follows that $CM_{j,j}$ is invertible for each
$j \in 2, 4, \ldots, 2\kappa$ and thus $CM_{even}(A, \lambda)$ is invertible too.

Let $n_{even} = n_{\lambda,2} + n_{\lambda,4} + \ldots + n_{\lambda,2\kappa}$. The numerical range

$$NR_{even}(A, \lambda) = \{\langle CM_{even}(A, \lambda)x, x\rangle \quad | x \in \mathbb{C}^{n_{even}}, x \neq 0\}$$

of $CM_{even}(A, \lambda)$ is called the characteristic *even numerical range* of A at λ.

Lemma 2.4.2 *Assume that $n_{even} = n_{\lambda,2} + n_{\lambda,4} + \ldots + n_{\lambda,2\kappa}$ as before. Suppose $x_1, x_2, \ldots, x_{n_{even}}$ denote the vectors in the set*

$$\{y_{2,1}, \ldots, y_{2,n_{\lambda,2}}\} \cup \{y_{4,1}, \ldots, y_{4,n_{\lambda,4}}\} \cup \ldots \cup \{y_{2\kappa,1}, \ldots, y_{2\kappa,n_{\lambda,2\kappa}}\}.$$

For $\xi = (\xi_1, \xi_2, \ldots, \xi_{n_{even}}) \in \mathbb{C}^{n_{even}}$, let $x = \sum_{j=1}^{n_{even}} \xi_j x_j$. Let $CM_{even}(A, \lambda)$ be defined as before with respect to the above vectors. Then

$$\langle CM_{even}(A, \lambda)\xi, \xi \rangle = [(A - \lambda)x, x].$$

Proof. On the one hand, let $n_{\lambda,0} = 0$. The vector $CM_{even}(A, \lambda)\xi$ consists of κ blocks, the first block being

$$\sum_{r=1}^{\kappa} \begin{bmatrix} \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2,1}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \\ \vdots \\ \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2,n_{\lambda,2}}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \end{bmatrix}.$$

The m^{th} block in the vector is given by

$$\sum_{r=1}^{\kappa} \begin{bmatrix} \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2m,1}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \\ \vdots \\ \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2m,n_{\lambda,2m}}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \end{bmatrix}.$$

So,

$$\begin{aligned}
\langle CM_{even}(A, \lambda)\xi, \xi \rangle &= \overline{\xi}_1 \sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2,1}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \\
&+ \ \ldots \\
&+ \ \overline{\xi}_{n_{\lambda,2}+1} \sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{4,1}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \\
&+ \ \ldots \\
&+ \ \overline{\xi}_{(\sum_{p=1}^{\kappa} n_{\lambda,2p-2})+1} \sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} [(A - \lambda)y_{2r,s}, y_{2\kappa,1}]\xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} \\
&+ \ \ldots \\
&= \sum_{l,r=1}^{\kappa} \sum_{t=1}^{n_{\lambda,2l}} \sum_{s=1}^{n_{\lambda,2r}} \overline{\xi}_{(\sum_{q=1}^{l} n_{\lambda,2q-2})+t} \xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} [(A - \lambda)y_{2r,s}, y_{2l,n_{\lambda,2l}}]
\end{aligned}$$

Now let $x_1, x_2, \ldots, x_{n_{even}}$ denote the vectors in the set

$$\{y_{2,1}, \ldots, y_{2,n_{\lambda,2}}\} \cup \{y_{4,1}, \ldots, y_{4,n_{\lambda,4}}\} \cup \ldots \cup \{y_{2\kappa,1}, \ldots, y_{2\kappa,n_{\lambda,2\kappa}}\}.$$

Put $x = \xi_1 x_1 + \ldots + \xi_{n_{even}} x_{n_{even}}$, then on the other hand we have

$$\begin{aligned}
[(A - \lambda)x, x] &= [(A - \lambda) \sum_{j=1}^{n_{even}} \xi_j x_j, \sum_{i=1}^{n_{even}} \xi_i x_i] \\
&= \sum_{i=1}^{n_{even}} [(A - \lambda) \sum_{j=1}^{n_{even}} \xi_j x_j, \xi_i x_i],
\end{aligned}$$

where

$$\sum_{j=1}^{n_{even}} \xi_j x_j = \sum_{j=1}^{n_{\lambda,2}} \xi_j y_{2,j} + \sum_{j=1}^{n_{\lambda,4}} \xi_{n_{\lambda,2}+j} y_{4,j} + \ldots + \sum_{j=1}^{n_{\lambda,2\kappa}} \xi_{(\sum_{p=1}^{\kappa} n_{\lambda,2p-2})+j} y_{2\kappa,j}.$$

Therefore,

$$\begin{aligned}
\sum_{i=1}^{n_{even}} [(A-\lambda) \sum_{j=1}^{n_{even}} \xi_j x_j, \xi_i x_i] &= \overline{\xi}_1 [\sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} \xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} y_{2r,s}, (A-\lambda)^{[*]} y_{2,1}] \\
&+ \ldots \\
&+ \overline{\xi}_{n_{\lambda,2}+1} [\sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} \xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} y_{2r,s}, (A-\lambda)^{[*]} y_{4,1}] \\
&+ \ldots \\
&+ \overline{\xi}_{(\sum_{p=1}^{\kappa} n_{\lambda,2p-2})+1} [\sum_{r=1}^{\kappa} \sum_{s=1}^{n_{\lambda,2r}} \xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} y_{2r,s}, (A-\lambda)^{[*]} y_{2\kappa,1}] \\
&+ \ldots \\
&= \sum_{l,r=1}^{\kappa} \sum_{t=1}^{n_{\lambda,2l}} \sum_{s=1}^{n_{\lambda,2r}} \overline{\xi}_{(\sum_{q=1}^{l} n_{\lambda,2q-2})+t} \xi_{(\sum_{p=1}^{r} n_{\lambda,2p-2})+s} [(A-\lambda) y_{2r,s}, y_{2l,n_{\lambda,2l}}].
\end{aligned}$$

Thus, we have that

$$\langle CM_{even}(A,\lambda)\xi, \xi \rangle = [(A-\lambda)x, x].$$

□

By using Lemma 2.4.2 and formula (3.18) of [77] we are now ready to obtain a relation between the respective even numerical range conditions. Let $x'_1, x'_2, \ldots, x'_{n_{even}}$ denote the vectors in the set

$$\{y'_{2,1}, \ldots, y'_{2,n_{\lambda,2}}\} \cup \{y'_{4,1}, \ldots, y'_{4,n_{\lambda,4}}\} \cup \ldots \cup \{y'_{2\kappa,1}, \ldots, y'_{2\kappa,n_{\lambda,2\kappa}}\}.$$

Let $\xi = (\xi_1, \xi_2, \ldots, \xi_{n_{even}}) \in \mathbb{C}^{n_{even}}$ be arbitrary. Then it follows from [77], p.61 that

$$\langle CM_{even}(iA, i\lambda)\xi, \xi \rangle = [(iA - i\lambda)x', x'],$$

such that $x' = \sum_{j=1}^{n_{even}} \xi_j x'_j$. Then

$$\begin{aligned}
x' &= (\xi_1 i^{\frac{6}{2}} y_{2,1} + \ldots + \xi_{n_{\lambda,2}} i^{\frac{6}{2}} y_{2,n_{\lambda,2}}) \\
&+ (\xi_{n_{\lambda,2}+1} i^{\frac{12}{2}} y_{4,1} + \ldots + \xi_{n_{\lambda,2}+n_{\lambda,4}} i^{\frac{12}{2}} y_{4,n_{\lambda,4}}) \\
&+ \ldots \\
&+ (\xi_{\sum_{p=1}^{\kappa} n_{\lambda,2p-2})+1} i^{3\kappa} y_{2\kappa,1} + \ldots + \xi_{n_{even}} i^{3\kappa} y_{2\kappa,n_{\lambda,2\kappa}}).
\end{aligned}$$

Now, let $x_1, x_2, \ldots, x_{n_{even}}$ denote the vectors

$$\{y_{2,1}, \ldots, y_{2,n_{\lambda,2}}\} \cup \{y_{4,1}, \ldots, y_{4,n_{\lambda,4}}\} \cup \ldots \cup \{y_{2\kappa,1}, \ldots, y_{2\kappa,n_{\lambda,2\kappa}}\}$$

and let

$$\mu_j = \begin{cases} (-i)\xi_j & \text{for} \quad j = 1, 2, \ldots, n_{\lambda,2} \\ (-1)\xi_j & \text{for} \quad j = n_{\lambda,2}+1, \ldots, n_{\lambda,2}+n_{\lambda,4} \\ (i)\xi_j & \text{for} \quad j = n_{\lambda,2}+n_{\lambda,4}+1, \ldots, n_{\lambda,2}+n_{\lambda,4}+n_{\lambda,6} \\ \vdots \end{cases}$$

Then $\mu = (\mu_1, \mu_2, \ldots, \mu_{n_{even}}) \in \mathbb{C}^{n_{even}}$ and

$$x' = \mu_1 x_1 + \mu_2 x_2 + \ldots + \mu_{n_{even}} x_{n_{even}}.$$

From Lemma 2.4.2 it now follows that

$$[(A-\lambda)x', x'] = \langle CM_{even}(A,\lambda)\mu, \mu\rangle$$

holds. Hence from [77], formula (3.18) we get that

$$\begin{aligned} \langle CM'_{even}(iA, i\lambda)\xi, \xi\rangle &= i[(A-\lambda)x', x'] \\ &= i\langle CM_{even}(A,\lambda)\mu, \mu\rangle. \end{aligned}$$

Therefore we can conclude that

$$0 \notin NR'_{even}(iA, i\lambda) = \{\langle CM'_{even}(iA, i\lambda)\xi, \xi\rangle | 0 \neq \xi \in \mathbb{C}^{n_{even}}\}$$

if and only if

$$0 \notin NR_{even}(A,\lambda) = \{\langle CM_{even}(A,\lambda)\mu, \mu\rangle | 0 \neq \mu \in \mathbb{C}^{n_{even}}\}.$$

Definition 2.4.3 *Let A be H-positive real, then the pair (A, H) is said to satisfy the even numerical range condition if $0 \notin NR_{even}(A,\lambda)$ for all $\lambda \in i\mathbb{R} \cap \sigma(A)$.*

Thus, from the above arguments it follows that (A, H) has the numerical range condition if and only if (iA, H) has the numerical range condition for the dissipative operators in [77].

Definition 2.4.4 *We say that a pair (A, H) satisfies the numerical range condition if it satisfies both the odd numerical range condition and the even numerical range condition.*

Example 2.4.5

Let

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

Now $HA + A^*H \geq 0$, hence A is H-positive real. It is clear that $\sigma(A) = \{0\}$, thus $x_1 = e_1$ and $z_1 = e_3$ are eigenvectors of A corresponding to its eigenvalue $\lambda = 0$. A simple calculation shows that a Jordan basis for $\mathcal{R}(A, \{0\})$ is given by $\{e_1, e_2, e_3\}$, with Jordan chains of length one and two given by:

$$J(0,1) = \{e_3\} \quad \text{and} \quad J(0,2) = \{e_1, e_2\}.$$

Referring to our discussion above, in this case: For j odd ($j = 1$ only) we have that $x_{1,1} = e_3$ and $y_{1,1} = e_3$. Therefore

$$CM_1 = ([y_{1,1}, y_{1,1}]) = -1$$

and hence $CM_{odd}(A, \{0\}) = -1$. Now let $n_{odd} = n_{0,1} = 1$, then

$$NR_{odd}(A, \{0\}) = (-\infty, 0)$$

and $0 \notin NR_{odd}(A, \{0\})$. Thus the odd numerical range condition is satisfied.

For j even ($j = 2$ only) we have $x_{2,1} = y_{2,1} = e_2$. Now,

$$CM_{2,2} = ([(A - \lambda)y_{2,1}, y_{2,1}]) = -i$$

and hence $CM_{even}(A, \{0\}) = -i$. Let $n_{even} = n_{0,2} = 1$, then

$$NR_{even}(A, \{0\}) = \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) = 0, \mathrm{Im}(z) < 0\}$$

and $0 \notin NR_{even}(A, \{0\})$. Thus the even numerical range condition is satisfied.

Thus, the pair (A, H) satisfies the numerical range condition for H-positive real matrices.

Consider the matrix iA, then

$$H(iA) - (iA)^* H \geq 0.$$

Hence iA is H-dissipative. As before $\sigma(iA) = \{0\}$. A simple calculation shows that a Jordan basis for $\mathcal{R}(iA, \{0\})$ is given by $\{e_1, -ie_2, e_3\}$, with Jordan chains of length one and two given by:

$$J'(0, 1) = \{e_3\} \quad \text{and} \quad J'(0, 2) = \{e_1, -ie_2\}.$$

For j odd ($j = 1$ only)

$$CM_1' = \left([y_{1,1}', y_{1,1}']\right) = -1$$

and hence $CM_{odd}'(iA, \{0\}) = -1$. Now let $n_{odd} = n_{0,1} = 1$, then

$$NR_{odd}'(iA, \{0\}) = (-\infty, 0) = NR_{odd}(A, \{0\})$$

and $0 \notin NR_{odd}'(iA, \{0\})$. Thus the odd numerical range condition is satisfied for dissipative matrices.

For j even ($j = 2$ only) we have $x_{2,1}' = y_{2,1}' = e_2$. Now,

$$CM_{2,2}' = ([(iA - i\lambda)y_{2,1}', y_{2,1}']) = -i^2 = 1$$

and hence $CM_{even}'(iA, \{0\}) = 1$. Let $n_{even} = n_{0,2} = 1$, then

$$NR_{even}'(iA, \{0\}) = (0, \infty) = iNR_{even}(A, \{0\})$$

and $0 \notin NR_{even}(A, \{0\})$. Thus the even numerical range condition is satisfied for dissipative matrices.

Thus, the pair (iA, H) satisfies the numerical range condition for dissipative matrices. □

Observe in the next example that the numerical range condition is not stable, i.e., it may fail to hold for H-positive real matrices arbitrarily close to A.

Example 2.4.6

Let

$$A_\epsilon = \begin{bmatrix} 0 & 1 & 0 \\ \epsilon^2 & 0 & 0 \\ 0 & 0 & -\epsilon \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Now $HA_\epsilon + A_\epsilon^* H \geq 0$, hence A_ϵ is H-positive real. The eigenvalues of A_ϵ are found to be $\{\epsilon, -\epsilon\}$. The eigenvector of A_ϵ corresponding to $\lambda = \epsilon$ is (say) $x_1 = (1, \epsilon, 0)^T$ and the eigenvectors of A_ϵ corresponding to $\lambda = -\epsilon$ is (say) $y_1 = (1, -\epsilon, 0)^T$ and $z_1 = (0, 0, 1)^T$. There are no generalized eigenvectors associated with any eigenvalues of the matrix A_ϵ. A simple calculation shows that a Jordan basis for $\mathcal{R}(A_\epsilon, \{-\epsilon\})$ is given by $\{y_1, z_1\}$, with Jordan chains of length one given by:

$$J(-\epsilon, 1) = \{y_1, z_1\}.$$

Thus, $n_{-\epsilon,1} = 2$. For j odd ($j = 1$ only) we have that $x_{1,1} = y_1$ and $x_{1,2} = z_1$. Hence, $y_{1,1} = y_1$ and $y_{1,2} = z_1$. Therefore

$$CM_1 = \begin{bmatrix} [y_{1,1}, y_{1,1}] & [y_{1,2}, y_{1,1}] \\ [y_{1,1}, y_{1,2}] & [y_{1,2}, y_{1,2}] \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and hence

$$CM_{odd}(A_\epsilon, \{-\epsilon\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now if $x \in \mathbb{C}^{n_{odd}}$is a nonzero vector, where $n_{odd} = n_{-\epsilon,1} = 2$, one finds that $0 \in NR_{odd}(A_\epsilon, \{-\epsilon\})$. Hence the odd numerical range condition is not satisfied for the pair (A_ϵ, H). Therefore the numerical range condition for the pair (A_ϵ, H) is not satisfied. □

2.5 Uniqueness and stability

Let $\mathcal{M}$ and $\mathcal{N}$ be subspaces of $\mathbb{C}^n$ and let $\theta(\mathcal{M}, \mathcal{N}) = \|P_\mathcal{M} - P_\mathcal{N}\|$ where $P_\mathcal{M}$ and $P_\mathcal{N}$ are the unique orthogonal projections of $\mathcal{M}$ and $\mathcal{N}$. Then it is clear that θ is a metric on the set $S(\mathbb{C}^n)$ of all subspaces of $\mathbb{C}^n$.

Definition 2.5.1 *Let $A \in \mathbb{A}$, where $\mathbb{A}$ denotes the class of all $n \times n$ H-positive real matrices. We call an A-invariant maximal H-nonnegative (respectively, H-nonpositive) subspace $\mathcal{M}$,* $\mathbb{A}$-stable *if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every $B \in \mathbb{A}$ with $\|A - B\| < \delta$ there exists a H-nonnegative (respectively, H-nonpositive) B-invariant subspace $\mathcal{N}$, such that*

$$\theta(\mathcal{M}, \mathcal{N}) < \epsilon.$$

Definition 2.5.2 *Let $A \in \mathbb{D}$, where $\mathbb{D}$ denotes the class of all H-dissipative matrices. We call an A-invariant maximal H-nonnegative (respectively, H-nonpositive) subspace $\mathcal{M}$,* $\mathbb{D}$-stable *if for every $\epsilon > 0$ there exist a $\delta > 0$ such that for every $B \in \mathbb{D}$ with $\|A - B\| < \delta$ there exist a H-nonnegative (respectively, H-nonpositive) B-invariant subspace $\mathcal{N}$, such that*

$$\theta(\mathcal{M}, \mathcal{N}) < \epsilon.$$

In the theorem to follow, we make use of Theorem 2.1 in [63]. We therefore first quote this theorem.

Theorem 2.5.3 *The following statements are equivalent for given H-dissipative matrix A:*

(i) there exists a stable A-invariant maximal H-nonnegative subspace,

(ii) there exists a stable A-invariant maximal H-nonpositive subspace,

(iii) the numerical range condition, in the sense of [77]*, holds for the pair (A, H).*

(iv) there is a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}$, with $\sigma(A|_{\mathcal{M}})$ contained in the closed upper half plane,

(v) there is a unique A-invariant maximal H-nonpositive subspace $\mathcal{M}$, with $\sigma(A|_{\mathcal{M}})$ contained in the closed lower half plane.

In that case, there is a unique stable A-invariant maximal H-nonnegative subspace, being the one with $\sigma(A|_{\mathcal{M}})$ contained in the closed upper half plane, and there is a unique stable A-invariant maximal H-nonpositive subspace, being the one with $\sigma(A|_{\mathcal{M}})$ contained in the closed lower half plane.

Theorem 2.5.4 *The following statements are equivalent for a given H-positive real matrix A:*

(i) there exists an $\mathbb{A}$-stable A-invariant maximal H-nonnegative subspace, say $\mathcal{M}_+$,

(ii) there exists an $\mathbb{A}$-stable A-invariant maximal H-nonpositive subspace, say $\mathcal{M}_-$,

(iii) the numerical range condition holds for the pair (A, H),

(iv) there is a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, with $\sigma(A|_{\mathcal{M}_+})$ contained in the closed right half plane,

(v) there is a unique A-invariant maximal H-nonpositive subspace $\mathcal{M}_-$, with $\sigma(A|_{\mathcal{M}_-})$ contained in the closed left half plane.

Proof. First, note that the pair (A, H) where $A \in \mathbb{A}$ satisfies the numerical range condition of Definition 2.4.4 if and only if the pair (iA, H) satisfies the numerical range condition in the sense of [77].

We prove (i)$\Leftrightarrow$ (iii). Assume (i): That is, there exists a $\mathbb{A}$-stable A-invariant H-nonnegative subspace $\mathcal{M}_+$. Notice that since $A \in \mathbb{A}$, A is H-positive real implies that iA is H-dissipative and also that $\mathcal{M}_+$ is iA- invariant.

It follows from Definition 2.5.1 that if we let $\epsilon > 0$, there exists a $\delta > 0$ such that for all $B \in \mathbb{A}$ with $\|A - B\| < \delta$ there exists a H-nonnegative B-invariant subspace $\mathcal{N}_+$, such that $\theta(\mathcal{M}_+, \mathcal{N}_+) < \epsilon$. Let B_0 be any H-dissipative matrix such that $\|iA - B_0\| < \delta$. Since B_0 is H-dissipative it follows that iB_0 is H-positive real, hence $iB_0 \in \mathbb{A}$ and thus

$$\|A - (-iB_0)\| = \|iA - B_0\| < \delta.$$

Hence, there exists a $-iB_0$-invariant H-nonnegative subspace $\mathcal{N}_+$ such that $\theta(\mathcal{M}_+, \mathcal{N}_+) < \epsilon$. But $\mathcal{N}_+$ is also B_0-invariant, hence $\mathcal{M}_+$ is $\mathbb{D}$-stable with respect to the pair (iA, H). Therefore, (i) holds in Theorem 2.5.3 and hence also (iii) in Theorem 2.5.3 for the pair (iA, H).

Conversely, assume that (iii) holds, i.e. the pair (iA, H) satisfies the numerical range condition. By Theorem 2.5.3 there exists a $\mathbb{D}$-stable iA-invariant maximal H-nonnegative subspace $\mathcal{M}_+$. The subspace $\mathcal{M}_+$ is clearly also A-invariant. We now verify that it is also $\mathbb{A}$-stable: Let $\epsilon > 0$ be given. Then there exists a $\delta > 0$ such that for every $B \in \mathbb{D}$ with $\|iA - B\| < \delta$, there exists a nonnegative B-invariant subspace $\mathcal{N}_+$ such that $\theta(\mathcal{M}_+, \mathcal{N}_+) < \epsilon$ (cf. Definition 2.5.2). Now, if we let $B_0 \in \mathbb{A}$ with $\|A - B_0\| < \delta$, then $iB_0 \in \mathbb{D}$ and also

$$\|iA - iB_0\| = \|A - B_0\| < \delta.$$

Therefore there exists a iB_0-invariant H-nonnegative subspace $\mathcal{N}_{+0}$ such that $\theta(\mathcal{M}_+, \mathcal{N}_{+0}) < \epsilon$. Since $\mathcal{N}_{+0}$ is also B_0-invariant, it follows that $\mathcal{M}_+$ is $\mathbb{A}$-stable with respect to the pair (A, H). This proves (i)$\Leftrightarrow$ (iii).

The equivalence of (ii) and (iii) follows similarly.

We now prove (iii)$\Leftrightarrow$ (iv): Assume (iv) holds. There exists a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, with $\sigma(A|_{\mathcal{M}_+})$ contained in the closed right half plane. Take note that, $A \in \mathbb{A}$ implies that iA is H-dissipative and that $\mathcal{M}_+$ is also iA-invariant. From Theorem 2.5.3, $\mathcal{M}_+$ with $\sigma(iA|_{\mathcal{M}_+})$ contained in the closed upper half plane is unique for the pair (iA, H). Thus, (iii) holds in Theorem 2.5.3 and hence (iii) holds for the pair (iA, H).

Conversely, let the pair (iA, H) satisfy the numerical range condition. Take note, iA is H-dissipative and $\mathcal{M}_+$ is iA-invariant. From Theorem 2.5.3 it follows that there exists a unique iA-invariant maximal H- nonnegative subspace $\mathcal{M}_+$, with $\sigma(iA|_{\mathcal{M}_+})$ contained in the closed upper half plane. Furthermore, $\mathcal{M}_+$ is also A-invariant and $\sigma(iA|_{\mathcal{M}_+}) \subseteq$ "closed upper half plane" implies $\sigma(A|_{\mathcal{M}_+}) \subseteq$ "closed right half plane". Thus, there is a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}_+$, with $\sigma(A|_{\mathcal{M}_+})$ contained in the closed right half plane.

The equivalence of (iii) and (v) follows similarly. □

2.5.1 The real case

It is now clear (from the previous theorem and from the work done by [77]) to see whether or not there exist unique complex (real) A-invariant maximal H-nonnegative (nonpositive) subspaces, respectively. The question now arises whether there is any relation between the existence of the respectively complex and real subspaces. The following lemma gives the answer to this question.

Lemma 2.5.5 *The following statements are equivalent for a given real H-positive real matrix A.*

(a) There exists a unique complex A-invariant maximal H-nonnegative subspace $\mathcal{M}$, such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{right}$.

(b) There exists a unique real A-invariant maximal H-nonnegative subspace $\mathcal{M}$, such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{right}$.

(c) There exists a unique complex A-invariant maximal H-nonpositive subspace $\mathcal{M}$, such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{left}$.

(d) There exists a unique real A-invariant maximal H-nonpositive subspace $\mathcal{M}$, such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{left}$.

Proof. We consider A acting on $\mathbb{C}^n$, so $A : \mathbb{C}^n \to \mathbb{C}^n$, and, we define an indefinite inner product on $\mathbb{C}^n$ by $\langle Hx, y\rangle$, where $x, y \in \mathbb{C}^n$, and $\langle ., .\rangle$ denotes the standard inner product on $\mathbb{C}^n$.

Assume (a), i.e. there exists a complex A-invariant maximal H-nonnegative subspace $\mathcal{M}$, with $\sigma(A|_M) \subseteq \overline{\mathbb{C}}_{right}$, such that $\langle Hx, x\rangle \geq 0$ for all $x \in \mathcal{M}$.

By the construction at the end of Section 3.5 there is a real basis for $\mathcal{M}$, compare formula (2.15). So, let $\{x_1, x_2, \ldots, x_m\}$ be a real basis in $\mathcal{M}$, and let $\mathcal{N}$ be the linear span over $\mathbb{R}$, i.e.

$$\mathcal{N} = \{x = \sum_{j=1}^{m} \alpha_j x_j | \alpha_j \in \mathbb{R}\}.$$

Observe that $\mathcal{M} = \mathcal{N} + i\mathcal{N}$. Let $x \in \mathcal{N}$, then

$$\begin{aligned} x = \overline{x} &\Leftrightarrow \sum_{j=1}^{m} \overline{\alpha}_j x_j = \sum_{j=1}^{m} \alpha_j x_j \\ &\Leftrightarrow \overline{\alpha}_j = \alpha_j, \end{aligned}$$

for all j. Now $Ax = A\overline{x} = \overline{Ax} \in \mathcal{M}$, since $\mathcal{M}$ is A-invariant. But $x \in \mathcal{N}$ if and only if $x \in \mathcal{M}$ and $x = \overline{x}$, thus $Ax \in \mathcal{N}$. Hence $\mathcal{N}$ is A-invariant. Also, since $\mathcal{N}$ is considered as a subspace over $\mathbb{R}$ and has the same basis as $\mathcal{M}$ over $\mathbb{C}$, it follows that $\dim \mathcal{N} = \dim \mathcal{M} = m$. Furthermore, $\mathcal{N}$ is H-nonnegative since $\mathcal{N} \subset \mathcal{M}$, thus if $x \in \mathcal{N} \subset \mathcal{M}$ it follows that $\langle Hx, x\rangle \geq 0$. Hence we have that $\mathcal{N}$ is A-invariant and H-nonnegative. The subspace $\mathcal{N}$ satisfies $\sigma(A|_{\mathcal{N}}) = \sigma(A|_{\mathcal{M}}) \subset \overline{\mathbb{C}}_{right}$: Let $\lambda \in \sigma(A|_{\mathcal{M}})$, then there exist a $0 \neq x \in \mathcal{M}$, such that $Ax = \lambda x$. Since $\mathcal{M} = \mathcal{N} + i\mathcal{N}$, there exist $x_1, x_2 \in \mathcal{N}$ such that $x = x_1 + ix_2$ (where either $x_1 \neq 0$ or $x_2 \neq 0$ or both $x_1, x_2 \neq 0$). Now, $Ax = \lambda x \Leftrightarrow Ax_1 + iAx_2 = \lambda x_1 + i\lambda x_2$, i.e., $Ax_1 = \lambda x_1$ and $Ax_2 = \lambda x_2$. Thus, $\sigma(A|_{\mathcal{M}}) \subseteq \sigma(A|_{\mathcal{N}})$. The inclusion $\sigma(A|_{\mathcal{N}}) \subseteq \sigma(A|_{\mathcal{M}})$ is obvious.

We now prove the maximality of $\mathcal{N}$. To see this, let $\tilde{\mathcal{N}}$ be a real H-nonnegative subspace such that $\sigma(A|_{\tilde{\mathcal{N}}}) \subseteq \overline{\mathbb{C}}_{right}$ and $\mathcal{N} \subseteq \tilde{\mathcal{N}}$ and $A(\tilde{\mathcal{N}}) \subseteq \tilde{\mathcal{N}}$. Let $\tilde{\mathcal{M}} = \tilde{\mathcal{N}} + i\tilde{\mathcal{N}}$, then $\mathcal{M} \subseteq \tilde{\mathcal{M}}$. Now $A(\tilde{\mathcal{M}}) = A(\tilde{\mathcal{N}}) + iA(\tilde{\mathcal{N}}) \subseteq \tilde{\mathcal{N}} + i\tilde{\mathcal{N}} = \tilde{\mathcal{M}}$, i.e. $\tilde{\mathcal{M}}$ is A-invariant. Let $x \in \tilde{\mathcal{M}}$ then $x = x_1 + ix_2$ with $x_1, x_2 \in \tilde{\mathcal{N}}$. Then

$$\begin{aligned} \langle H(x_1 + ix_2), x_1 + ix_2\rangle &= \langle Hx_1, x_1\rangle + i\langle Hx_2, x_1\rangle - i\langle Hx_1, x_2\rangle + \langle Hx_2, x_2\rangle \\ &= \langle Hx_1, x_1\rangle + \langle Hx_2, x_2\rangle \\ &\geq 0, \end{aligned}$$

which proves $\tilde{\mathcal{M}}$ is H-nonnegative. As before $\sigma(A|_{\tilde{\mathcal{M}}}) = \sigma(A|_{\tilde{\mathcal{N}}}) \subseteq \overline{\mathbb{C}}_{right}$. The maximality of $\mathcal{M}$ implies that $\mathcal{M} = \tilde{\mathcal{M}}$. Thus, $\mathcal{N} = \tilde{\mathcal{N}}$ and therefore $\mathcal{N}$ is maximal.

To prove the uniqueness, let $\mathcal{N}_1$ and $\mathcal{N}_2$ be two real A-invariant maximal H-nonnegative subspaces such that $\sigma(A|_{N_j}) \subseteq \overline{\mathbb{C}}_{right}$ for $j = 1, 2$. Set $M_j = \{x + iy | x, y \in \mathcal{N}_j\}$. As before $\mathcal{M}_j$ is A-invariant and H-nonnegative such that $\sigma(A|_{\mathcal{M}_j}) \subseteq \overline{\mathbb{C}}_{right}$. We prove that $\mathcal{M}_j$ is maximal with these properties: Suppose not. Let $\tilde{\mathcal{M}}_j$ have the same properties where $\mathcal{M}_j \subset \tilde{\mathcal{M}}_j$. As before, let $\tilde{\mathcal{M}}_j = \tilde{\mathcal{N}}_j + i\tilde{\mathcal{N}}_j$ with $\tilde{\mathcal{N}}_j$ being a real A-invariant, maximal and H-nonnegative subspace. Then $\mathcal{N}_j \subset \tilde{\mathcal{N}}_j$ for $j = 1, 2$. The maximality of $\mathcal{N}_j$ with these properties then implies that $\mathcal{N}_j = \tilde{\mathcal{N}}_j$, i.e. $\mathcal{M}_j = \tilde{\mathcal{M}}_j$. By the uniqueness property it follows that $\mathcal{M}_1 = \mathcal{M}_2$. Thus, $\mathcal{N}_1 = \mathcal{N}_2$.

Therefore, there exists a unique real A-invariant maximal H-nonnegative subspace $\mathcal{M}$ such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{right}$. Thus we have proved that (a) implies (b).

Assume (b). Let $\mathcal{N}$ be the unique real A-invariant maximal H-nonnegative subspace such that $\sigma(A|_{\mathcal{N}}) \subseteq \overline{\mathbb{C}}_{right}$. Let $\mathcal{M} = \{x + iy | x, y \in \mathcal{N}\}$, then clearly $\mathcal{M}$ is A-invariant and $\mathcal{M}$ is H-nonnegative. Also, as before $\sigma(A|_{\mathcal{M}}) = \sigma(A|_{\mathcal{N}}) \subseteq \overline{\mathbb{C}}_{right}$. The proof of the maximality is similar to the proof of the maximality in (a) implies (b).

The uniqueness argument also follows similarly as above. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two complex A-invariant maximal H-nonnegative subspaces such that $\sigma(A|_{\mathcal{M}_j}) \subseteq \overline{\mathbb{C}}_{right}$ for $j = 1, 2$. Let $\mathcal{N}_j$ for $j = 1, 2$ be the real subspaces such that $\mathcal{M}_j = \mathcal{N}_j + i\mathcal{N}_j$ as before. Recall that $\mathcal{N}_j$ is A-invariant and maximal H-nonnegative such that $\sigma(A|_{\mathcal{N}_j}) = \sigma(A|_{\mathcal{M}_j}) \subseteq \overline{\mathbb{C}}_{right}$. By the uniqueness assumption in (b), it follows that $\mathcal{N}_1 = \mathcal{N}_2$. Hence $\mathcal{M}_1 = \mathcal{M}_2$.

Thus, there exists a unique complex A-invariant maximal H-nonnegative subspace $\mathcal{M}$ such that $\sigma(A|_{\mathcal{M}}) \subseteq \overline{\mathbb{C}}_{right}$.

The proof of (c) equivalent to (d) follows similarly. Since (a) is equivalent to (c) by Theorem 2.5.4 all the statements of the Lemma are equivalent. □

We want to conclude this section by giving some general examples that appear in the real case.

Example 2.5.6

Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, and $H = \begin{bmatrix} 0 & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$. Assume that $HA + A^T H \geq 0$. A direct computation gives

$$D := HA + A^T H = \begin{bmatrix} 0 & 0 \\ 0 & 2h_{12} \end{bmatrix}.$$

Now for A to be H-positive real, the matrix D needs to be positive semidefinite. In particular, both diagonal entries of D need to be greater than or equal to zero. So $h_{12} \geq 0$, and as H is assumed to be invertible we actually have $h_{12} \neq 0$.

Note that this is different from the case of H-skew symmetric matrices, for which we have $HA + A^T H = 0$. In this case, even sized Jordan blocks necessarily come in coupled pairs, see the canonical form in [40], also [66], Section 6.

Example 2.5.7

Let $A = \text{diag } J_2(0) \oplus J_2(0)$, $H = \begin{bmatrix} 0 & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix}$ and assume that $HA + A^T H \geq 0$. A direct computation gives

$$HA + A^T H = \begin{bmatrix} 0 & 0 \\ 0 & H_{12} + H_{12}^T \end{bmatrix}.$$

Thus, we need $H_{12} + H_{12}^T \geq 0$. Setting $H_{12} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have $H_{12} + H_{12}^T = \begin{bmatrix} 2a & b+c \\ b+c & 2d \end{bmatrix}$. So, we need $a \geq 0, d \geq 0$ and $4ad - (b+c)^2 \geq 0$. Note that in this example it is possible that $HA + A^T H = 0$, just take $a = 0, d = 0$ and $b = -c$.

Observe that the situation in Example 2.5.7 is actually like the pure imaginary eigenvalues, where we have a block of order two at $i\gamma$ and at $-i\gamma$, only now we set $\gamma = 0$.

Remark. Even sized Jordan blocks corresponding to the zero eigenvalue come in two flavors: there are some that come uncoupled as in Example 2.5.6, and there are some that come in coupled pairs, as in the canonical form for H-skew symmetric matrices, compare Example 2.5.7 for the case where $H_{12} + H_{12}^T = 0$.

Example 2.5.8

As a more general example, consider a real $2n \times 2n$ matrix A in the following special Jordan form and with eigenvalue $\lambda = 0$ (of multiplicity $2n$):

$$A = \begin{bmatrix} 0 & 1 & & & & & & \\ 0 & 0 & & & & & & \\ & & 0 & 1 & & & & \\ & & 0 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & 1 \\ & & & & & 0 & 0 \end{bmatrix}.$$

Thus, from A we see that there are n Jordan chains of length 2. Suppose the corresponding real matrix H has the following form:

$$H = \left[\begin{array}{cc|cc|cc|cc|c|cc} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & * & a_{1,1} & * & a_{1,2} & * & a_{1,3} & * & \cdots & a_{1,n-1} & * \\ \hline 0 & a_{1,1} & 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & * & 1 & * & a_{2,1} & * & a_{2,2} & * & \cdots & a_{2,n-2} & * \\ \hline 0 & a_{1,2} & 0 & a_{2,1} & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & * & 0 & * & 1 & * & a_{3,1} & * & \cdots & a_{3,n-3} & * \\ \hline \vdots & & \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ \hline 0 & a_{1,n-1} & 0 & a_{2,n-2} & 0 & a_{3,n-3} & 0 & 0 & \cdots & 0 & 1 \\ 0 & * & 0 & * & 0 & * & 0 & * & \cdots & 1 & * \end{array}\right].$$

Let A and H be as above and let $B = HA + A^T H$. An easy calculation shows that:

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 2 & 0 & a_{1,1} & 0 & a_{1,2} & \dots & 0 & a_{1,n-1} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & a_{1,1} & 0 & 2 & 0 & a_{2,1} & \dots & 0 & a_{2,n-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & a_{1,2} & 0 & a_{2,1} & 0 & 2 & \dots & 0 & a_{3,n-3} \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & a_{1,n-1} & 0 & a_{2,n-2} & 0 & a_{3,n-3} & \dots & 0 & 2 \end{bmatrix}.$$

For the given matrix A, one easily verifies that

$$CM_{even}(A, \{0\}) = CM_{2,2} = \begin{bmatrix} 1 & a_{1,1} & a_{1,2} & \dots & a_{1,n-1} \\ 0 & 1 & a_{2,1} & \dots & a_{2,n-2} \\ 0 & 0 & 1 & \dots & a_{3,n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Also, if we let $C = CM_{even}(A, \{0\}) + CM_{even}(A, \{0\})^T$, then

$$C = \begin{bmatrix} 2 & a_{1,1} & a_{1,2} & \dots & a_{1,n-1} \\ a_{1,1} & 2 & a_{2,1} & \dots & a_{2,n-2} \\ a_{1,2} & a_{2,1} & 2 & \dots & a_{3,n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1,n-1} & a_{2,n-2} & a_{3,n-3} & \dots & 2 \end{bmatrix}.$$

Comparing the $n \times n$ matrix C with the $2n \times 2n$ matrix B, we see that if $x = (\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n$ and $\tilde{x} = (0, \alpha_1, 0, \alpha_2, 0, \dots, 0, \alpha_n) \in \mathbb{C}^{2n}$, then $\langle Cx, x\rangle = \langle B\tilde{x}, \tilde{x}\rangle$. Thus if $B \geq 0$, then $C \geq 0$. On the other hand, if $x = (\beta_1, \dots, \beta_{2n}) \in \mathbb{C}^{2n}$ and $\tilde{x} = (0, \beta_2, \dots, 0, \beta_{2n}) \in \mathbb{C}^{2n}$, then $\langle Bx, x\rangle = \langle B\tilde{x}, \tilde{x}\rangle = \langle Cy, y\rangle$, where $y = (\beta_2, \beta_4, \dots, \beta_{2n}) \in \mathbb{C}^n$. Thus, if $C \geq 0$, then $B \geq 0$. Hence $B \geq 0 \Leftrightarrow C \geq 0$. One verifies by direct calculation that $\langle Cx, x\rangle = 2\mathrm{Re}\langle CM_{even}(A, \{0\})x, x\rangle$ for all $x \in \mathbb{C}^n$. Thus, $C \geq 0$ if and only if $\mathrm{Re}\langle CM_{even}(A, \{0\})x, x\rangle \geq 0$. We conclude that A is H-positive real if and only if $CM_{even}(A, \{0\}) \geq 0$.

Suppose that the even numerical range condition holds at eigenvalue zero for all blocks of order two, i.e., $\mathrm{Re}\langle CM_{even}(A, \{0\})x, x\rangle > 0$ for $x \neq 0$. Then this implies that $\langle (CM_{even}(A, \{0\}) + CM_{even}(A, \{0\})^T)x, x\rangle > 0$ for $x \neq 0$. But $C = CM_{even}(A, \{0\}) + CM_{even}(A, \{0\})^T$ so $\langle Cx, x\rangle > 0$ for every $x \neq 0$. Since $C = C^T$ this implies that C has full rank.

In conclusion: The even numerical range condition holds if and only if C is positive definite.

Example 2.5.9

Let A and H have the above form which in general can be written as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad H = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix},$$

with $\sigma(A_{11}) = \sigma(A_{22}) = \{0\}$ and A is H-positive real. Furthermore we assume that $H^T = H$, with H_{11} positive semidefinite and H_{22} negative definite. (It is well known that every real symmetric matrix is similar to a diagonal matrix with its eigenvalues on the main diagonal, which may be ordered.)

For $\epsilon > 0$, consider

$$A_\epsilon = \begin{bmatrix} A_{11} + \epsilon I & A_{12} \\ 0 & A_{22} - \epsilon I \end{bmatrix},$$

i.e. we consider a small perturbation in A, then

$$HA_\epsilon + A_\epsilon^T H = HA + A^T H + 2\epsilon \begin{bmatrix} H_{11} & 0 \\ 0 & -H_{22} \end{bmatrix} \geq 0.$$

Since $H_{11} \geq 0$ and $H_{22} < 0$, we see that A_ϵ is H-positive real. Moreover $\sigma(A_\epsilon) \cap i\mathbb{R} = \emptyset$, because $\epsilon > 0$, hence $\sigma(A_{11} + \epsilon I) \subset \mathbb{C}_{right}$ and $\sigma(A_{22} - \epsilon I) \subset \mathbb{C}_{left}$. (This follows from the spectral theorem for matrix polynomials.) Hence, $CM_{even}(A_\epsilon, \{-\epsilon, \epsilon\}) > 0$ from above, which means that the even numerical range condition holds. Since we only have Jordan chains of even length, the odd numerical range condition is automatically satisfied. Thus, the numerical range condition holds for the pair (A_ϵ, H) and therefore by Theorem 2.5.4 there exists a unique A-invariant maximal H-nonnegative subspace $\mathcal{M}$, say.

Chapter 3

Simple forms and maximal invariant subspaces of H-expansive matrices

3.1 Introductory remarks

Let $H = H^T$ be an invertible $n \times n$ symmetric real matrix, and let A be an $n \times n$ real matrix satisfying $A^T H A \geq H$. Such a matrix is called H-*expansive* or i.e. $A^T H A - H \geq 0$.There exists a connection between H-expansive and H-positive real matrices through a suitable *Cayley transform*. To see this, let

$$B = (\eta + A)(\eta - A)^{-1},$$

with $\eta \notin \sigma(A)$ and $|\eta| = 1$. Note that $HB + B^*H = 2((\eta - A)^*)^{-1}(H - A^T H A)(\eta - A)^{-1} \leq 0$, that is $-B$ is H-positive real. (It can be somewhat confusing that this property is called H- dissipative in [24], see pages 61, 218 and following.) If $\eta \notin \mathbb{R}$, then B will be a complex matrix. Therefore the theory with respect to complex matrices can be applied and the B-invariant positive semidefinite and negative semidefinite subspaces can be constructed, see [19]. Take note that A can also be expressed in terms of B, that is

$$A = (B + I)^{-1}(B - I)\eta,$$

since $\eta \notin \sigma(A)$ implies $-1 \notin \sigma(B)$. Since B can be represented as a series in A and also A can be represented as a series of B, they have the same invariant subspaces.

For the real case, if $1 \notin \sigma(A)$ or $-1 \notin \sigma(A)$ then the same argument gives

$$B = (I + A)(I - A)^{-1} \quad \text{or} \quad B = (-I + A)(-I - A)^{-1},$$

respectively. Thus, if A is real so is B. Again, in this case $HB + B^T H \leq 0$. However, if both 1 and -1 are in $\sigma(A)$, then we cannot directly apply this argument. Of course we can apply Cayley transform $B = (\eta + A)(\eta - A)^{-1}$ with some $\eta \notin \sigma(A)$, but this B is now a complex matrix. Therefore, it is difficult to reconstruct real invariant semidefinite subspaces of A from the complex invariant semidefinite subspaces of B. We prefer to treat the real case directly. This also yields more insight in the structure of real H-expansive matrices and their invariant subspaces.

Let A and H be real. We know that the eigenvalues of A can lie on the real axis (single) or else they will appear as conjugate pairs in $\mathbb{C}$. The eigenvalues of interest are

those for which $|\lambda| = 1$. Indeed, consider $Ax = \lambda x$, with $0 \neq x \in \mathbb{C}^n, \lambda \in \mathbb{C}$ and $|\lambda| \neq 1$. Then

$$\begin{aligned} & \langle (A^T HA - H)x, x\rangle = \langle A^T HAx, x\rangle - \langle Hx, x\rangle \\ = \; & |\lambda|^2 \langle Hx, x\rangle - \langle Hx, x\rangle = (|\lambda|^2 - 1)\langle Hx, x\rangle. \end{aligned}$$

Thus, if $\langle (A^T HA - H)x, x\rangle \geq 0$, then either $|\lambda| > 1$ and $\langle Hx, x\rangle \geq 0$, or $|\lambda| < 1$ and $\langle Hx, x\rangle \leq 0$, or $|\lambda| = 1$.

If we let $\mathcal{M}_{int}$ be the span of all eigenvectors and generalized eigenvectors corresponding to the eigenvalues inside the unit disc, $\mathbb{D}$, then $\mathcal{M}_{int}$ is H-nonpositive. If $\mathcal{M}_{ext}$ is the span of all eigenvectors and all generalized eigenvectors corresponding to the eigenvalues outside $\mathbb{D}$, then $\mathcal{M}_{ext}$ is H-nonnegative, see [5]. Both $\mathcal{M}_{int}$ and $\mathcal{M}_{ext}$ are isomorphic to a real vector space.

Furthermore, if $A^T HA - H$ is strictly positive then either $|\lambda| > 1$ and $\langle Hx, x\rangle > 0$ or, $|\lambda| < 1$ and $\langle Hx, x\rangle < 0$. Hence $\mathcal{M}_{int}$ will then be maximal H-negative and $\mathcal{M}_{ext}$ will be maximal H-positive.

In this chapter we consider the explicit construction of A-invariant maximal H-nonnegative and A-invariant maximal H-nonpositive subspaces for a real matrix A which has eigenvalues λ with $|\lambda| = 1$. (Existence is known from e.g. [5].)

3.2 Preliminary lemmas

For a matrix X, we usually use the notation x_{ij} or $x_{i,j}$ to denote the (i, j)-scalar entry. Sometimes, however, it is convenient to denote the (i, j)-scalar entry of a block matrix X (when regarded as a usual matrix) by X_{ij}. Although it will mostly be clear from the context when we refer to either the (i, j)-scalar entry or the (i, j)-block entry, we agree to consistently use $X_{i,j}$ for the (i, j)-block of the block matrix X. Note the difference between X_{ij} and $X_{i,j}$.

Let $\lambda = \alpha \pm i\beta$ be an eigenvalue of A with $|\lambda| = 1$ (i.e. $\alpha^2 + \beta^2 = 1$) and assume $\beta \neq 0$. We therefore look at the case where the eigenvalues lie on the unit circle. If A is H-expansive we have $A^T HA - H \geq 0$. Let A be in real Jordan form (e.g., see [24], Section A2 and [75] Chapter 6). In this chapter, we consider H to be invertible and assume that $H = H^*$. As H is real, $H = H^T$.

We start with a remark that will be useful in the developements to follow.

Remark 3.2.1 *If $A^T HA - H \geq 0$ and for some i it holds that $(A^T HA - H)_{ii} = 0$, then $(A^T HA - H)_{ij} = 0$ for all j and $(A^T HA - H)_{ji} = 0$ for all j.*

Proof. If $X \geq 0$ and $\langle Xe_i, e_i\rangle = 0$, then $\langle Xe_i, e_j\rangle = 0$ for all j. Indeed, if $\langle Xe_i, e_i\rangle = 0$, then $\langle X^{1/2}e_i, X^{1/2}e_i\rangle = 0$ and hence $X^{1/2}e_i = 0$. This implies $Xe_i = 0$, and thus $\langle Xe_i, e_j\rangle = 0$ for all j. This means that the $(j, i)^{th}$-entry equals zero for all j. Since $X^T = X$, it follows that the $(i, j)^{th}$-entry of X equals zero for all j. Finally, take $X = A^T HA - H$. □

We want to find a simple form for H (A is already in a canonical form, namely the real Jordan normal form). We explain the procedure by a few simple examples.

Example 3.2.2

Let $A = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ where $\beta \neq 0$ and $\alpha^2 + \beta^2 = 1$. If $H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$ is invertible and A is H-expansive, then

$$\begin{aligned} A^T HA - H &= \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} - \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} \\ &= \begin{bmatrix} -\beta^2(h_{11} - h_{22}) - 2\alpha\beta h_{12} & \alpha\beta(h_{11} - h_{22}) - 2\beta^2 h_{12} \\ \alpha\beta(h_{11} - h_{22}) - 2\beta^2 h_{12} & \beta^2(h_{11} - h_{22}) + 2\alpha\beta h_{12} \end{bmatrix}. \end{aligned}$$

Since both entries on the main diagonal have to be nonnegative and since they have opposite signs, they have no other choice than to be equal to zero. Using Remark 3.2.1 it follows that $\gamma^T H\gamma - H = 0$. In particular, we have

$$\begin{aligned} \beta(h_{11} - h_{22}) + 2\alpha h_{12} &= 0 \\ \alpha(h_{11} - h_{22}) - 2\beta h_{12} &= 0, \end{aligned}$$

showing that $\begin{bmatrix} \beta & \alpha \\ \alpha & -\beta \end{bmatrix} \begin{bmatrix} h_{11} - h_{22} \\ 2h_{12} \end{bmatrix} = 0$. Since $\det \begin{bmatrix} \beta & \alpha \\ \alpha & -\beta \end{bmatrix} = -1 \neq 0$, we have that $h_{11} = h_{22}$ and $h_{12} = 0$. Hence, $H = h_{11} I$ for some h_{11}.

Example 3.2.3

Let

$$A = \begin{bmatrix} \gamma & I \\ 0 & \gamma \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{1,2}^T & H_{2,2} \end{bmatrix},$$

where $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ is the 2×2-block associated with the pair of eigenvalues $\lambda = \alpha \pm i\beta$. Each $H_{i,j}$ is a 2×2 block matrix. If A is H-expansive, we have

$$\begin{aligned} 0 &\leq A^T HA - H \\ &= \begin{bmatrix} \gamma^T H_{1,1}\gamma - H_{1,1} & \gamma^T H_{1,1} + \gamma^T H_{1,2}\gamma - H_{1,2} \\ H_{1,1}\gamma + \gamma^T H_{1,2}^T\gamma - H_{1,2}^T & H_{1,1} + H_{1,2}\gamma + \gamma^T H_{1,2}^T + \gamma^T H_{2,2}\gamma - H_{2,2} \end{bmatrix}. \end{aligned} \tag{3.1}$$

For $A^T HA - H$ to be positive semidefinite it is necessary that the blocks on the main diagonal should be greater than or equal to zero. So we must have

$$\gamma^T H_{1,1}\gamma - H_{1,1} \geq 0, \tag{3.2}$$

$$H_{1,1} + H_{1,2}\gamma + \gamma^T H_{1,2}^T + \gamma^T H_{2,2}\gamma - H_{2,2} \geq 0. \tag{3.3}$$

Note that (3.2) implies $H_{1,1} = h_{11} I$ by Example 3.2.2. For the $H_{1,2}$ and $H_{2,2}$ blocks we need some additional lemmas which will follow next. After these lemmas we come back to Example 3.2.3.

We first introduce notations and classes of matrices that we will use in the rest of this chapter. Let $K_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $K_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $H_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Furthermore, we let

$$\mathcal{E} = \{aI + bH_0 \,|\, a, b \in \mathbb{R}\}.$$

In particular $\gamma \in \mathcal{E}$. Using that $H_0^2 = -I$, the following lemma follows by straightforward calculation.

Lemma 3.2.4 *If $K, L \in \mathcal{E}$ then:*

$$\begin{aligned} KL &= LK \qquad (3.4)\\ KK^T &= K^TK = (k_{11}^2 + k_{12}^2)I. \qquad (3.5)\end{aligned}$$

The following lemma is the first of a series of lemmas that will be useful in the simplification of the matrix H into a form which we will use in the construction of the H-invariant subspaces.

Lemma 3.2.5 *Let $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ be such that $\beta \neq 0$ and $\alpha^2 + \beta^2 = 1$. If*

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix}$$

and $\gamma^T H\gamma - H \geq 0$, then $\gamma^T H\gamma - H = 0 \quad$ and $\quad H = h_{11}I$.

The proof of this lemma is actually given in Example 3.2.2. Along with Lemma 3.2.5 we state the following lemma:

Lemma 3.2.6 *Let $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ be such that $\beta \neq 0$ and $\alpha^2 + \beta^2 = 1$. If*

$$Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

is any non-symmetric real matrix with $\gamma^T Y\gamma - Y = 0$, then $Y \in \mathcal{E}$. Conversely, if $Y \in \mathcal{E}$ then $\gamma^T Y\gamma - Y = 0$.

Proof. As in the proof of Lemma 3.2.5 we have

$$\begin{aligned} 0 &= \gamma^T Y\gamma - Y \\ &= \begin{bmatrix} -\beta^2(y_{11} - y_{22}) - \alpha\beta(y_{12} + y_{21}) & \alpha\beta(y_{11} - y_{22}) - \beta^2(y_{12} + y_{21}) \\ \alpha\beta(y_{11} - y_{22}) - \beta^2(y_{12} + y_{21}) & \beta^2(y_{11} - y_{22}) + \alpha\beta(y_{12} + y_{21}) \end{bmatrix}. \end{aligned}$$

Since $\beta \neq 0$ we have, $\begin{bmatrix} \beta & \alpha \\ \alpha & -\beta \end{bmatrix} \begin{bmatrix} y_{11} - y_{22} \\ y_{12} + y_{21} \end{bmatrix} = 0$ and because $\det \begin{bmatrix} \beta & \alpha \\ \alpha & -\beta \end{bmatrix} \neq 0$, $y_{11} = y_{22}$ and $y_{21} = -y_{12}$. Hence, $Y = \begin{bmatrix} y_{11} & y_{12} \\ -y_{12} & y_{11} \end{bmatrix}$ or $Y = y_{11}I + y_{12}H_0 \in \mathcal{E}$.

To prove the converse, let $Y \in \mathcal{E}$, then

$$\begin{aligned} \gamma^T Y\gamma - Y &= \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} - \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \\ &= \begin{bmatrix} a & b \\ -b & a \end{bmatrix} - \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = 0. \end{aligned}$$

□

Since the matrix H in (3.1) satisfies $H = H^T$, we can apply Lemma 3.2.5 to the $(1,1)$-block of the matrix in (3.1) and obtain $\gamma^T H_{1,1}\gamma - H_{1,1} = 0$, furthermore $H_{1,1} = h_{11}I$. Applying Remark 3.2.1 to the $(1,2)$-block of the matrix in (3.1) we obtain

$$\gamma^T H_{1,1} + \gamma^T H_{1,2}\gamma - H_{1,2} = 0. \tag{3.6}$$

Lemma 3.2.7 *Let γ be defined as before, let X, Y be any 2×2 real matrices not necessarily symmetric and suppose that $\gamma^T X + \gamma^T Y\gamma - Y = 0$. Then $X = x_{11}K_0 + x_{12}K_1$. Furthermore if $X = x_{11}I + x_{12}H_0$, then it follows that $X = 0$ and $\gamma^T Y\gamma - Y = 0$ and thus $Y \in \mathcal{E}$.*

Proof. Let $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ and consider

$$\gamma^T X + \gamma^T Y\gamma - Y = 0. \tag{3.7}$$

Take the trace of equation (3.7), note that $\gamma\gamma^T = I$, and $tr(\gamma^T Y\gamma) = tr(\gamma\gamma^T Y)$. Then

$$\begin{aligned} 0 &= tr(\gamma^T X) + tr(\gamma^T Y\gamma - Y) \\ &= tr(\gamma^T X) + tr(\gamma\gamma^T Y) - tr(Y) \\ &= tr(\gamma^T X) \\ &= \alpha(x_{11} + x_{22}) + \beta(x_{12} - x_{21}). \end{aligned} \tag{3.8}$$

Multiply equation (3.7) by H_0 from the left and take the trace again; we have by using $\gamma H_0\gamma^T = H_0$ that

$$\begin{aligned} 0 &= tr(H_0\gamma^T X) + tr(H_0\gamma^T Y\gamma - H_0 Y) \\ &= tr(H_0\gamma^T X) + tr(\gamma H_0\gamma^T Y) - tr(H_0 Y) \\ &= tr(H_0\gamma^T X) \\ &= \beta(x_{11} + x_{22}) - \alpha(x_{12} - x_{21}). \end{aligned} \tag{3.9}$$

From equations (3.8) and (3.9) it follows that

$$\begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix} \begin{bmatrix} x_{11} + x_{22} \\ x_{12} - x_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and since $\det \begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix} \neq 0$, we have $x_{11} = -x_{22}$ and $x_{12} = x_{21}$. Hence $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & -x_{11} \end{bmatrix} = x_{11}K_0 + x_{12}K_1$. If also $X = x_{11}I + x_{12}H_0$, it is straightforward to see that X must be zero and hence $\gamma^T Y\gamma - Y = 0$ by (3.7). Therefore by Lemma 3.2.6, $Y \in \mathcal{E}$. □

Returning to (3.1), we see from Lemma 3.2.7 and Lemma 3.2.5 that $H_{1,1} = 0$. It then follows from Lemma 3.2.6 that $H_{1,2} \in \mathcal{E}$.

We continue with another lemma which will be useful in the sequel.

Lemma 3.2.8 *Let γ be defined as before and suppose $X = x_{11}I + x_{12}H_0$, $Y = y_{11}I + y_{12}H_0$ and let Z be any 2×2 real matrix, not necessarily symmetric, satisfying*

$$X\gamma + \gamma^T Y + \gamma^T Z\gamma - Z = 0. \tag{3.10}$$

Put $\Psi = -\gamma^2$. Then $X = \Psi^T Y$ and $Y = \Psi X$. Furthermore, $X\gamma + \gamma^T Y = 0$ and $\gamma^T Z\gamma - Z = 0$, hence $Z \in \mathcal{E}$. In particular if Z is symmetric, then $Z = cI$ for some constant c.

Proof. Take the trace of equation (3.10). As in the proof of Lemma 3.2.7, $tr(\gamma^T Z\gamma - Z) = 0$ and $tr(H_0\gamma^T Z\gamma - H_0 Z) = 0$. Thus we have

$$\begin{aligned} 0 &= tr(X\gamma) + tr(\gamma^T Y) \\ &= tr\left(\begin{bmatrix} \alpha x_{11} - \beta x_{12} & \beta x_{11} + \alpha x_{12} \\ -\alpha x_{12} - \beta x_{11} & -\beta x_{12} + \alpha x_{11} \end{bmatrix}\right) \\ &\quad + tr\left(\begin{bmatrix} \alpha y_{11} + \beta y_{12} & \alpha y_{12} - \beta y_{11}) \\ \beta y_{11} - \alpha y_{12} & \beta y_{12} + \alpha y_{11} \end{bmatrix}\right) \\ &= 2\alpha x_{11} - 2\beta x_{12} + 2\alpha y_{11} + 2\beta y_{12}. \end{aligned} \tag{3.11}$$

Pre-multiplying equation (3.10) by H_0 and taking the trace again we have

$$\begin{aligned} 0 &= tr(H_0 X\gamma) + tr(H_0\gamma^T Y) \\ &= tr\left(\begin{bmatrix} -\alpha x_{12} - \beta x_{11} & -\beta x_{12} + \alpha x_{11} \\ -\alpha x_{11} + \beta x_{12} & -\beta x_{11} - \alpha x_{12} \end{bmatrix}\right) \\ &\quad + tr\left(\begin{bmatrix} \beta y_{11} - \alpha y_{12} & \beta y_{12} + \alpha y_{11} \\ -\alpha y_{11} - \beta y_{12} & -\alpha y_{12} + \beta y_{11} \end{bmatrix}\right) \\ &= -2\alpha x_{12} - 2\beta x_{11} - 2\alpha y_{12} + 2\beta y_{11}. \end{aligned} \tag{3.12}$$

From equations (3.11) and (3.12) it follows that

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} -\alpha & -\beta \\ \beta & -\alpha \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix}.$$

Since $\det \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} = 1$ we have

$$\begin{aligned} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} &= \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}^{-1} \begin{bmatrix} -\alpha & -\beta \\ \beta & -\alpha \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \\ &= \begin{bmatrix} \beta^2 - \alpha^2 & -2\alpha\beta \\ 2\alpha\beta & \beta^2 - \alpha^2 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \\ &= \Psi \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix}. \end{aligned}$$

The second column of X is therefore given by

$$\begin{aligned}\begin{bmatrix} x_{12} \\ x_{11} \end{bmatrix} &= \begin{bmatrix} 2\alpha\beta & \beta^2-\alpha^2 \\ \beta^2-\alpha^2 & -2\alpha\beta \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \\ &= \begin{bmatrix} \beta^2-\alpha^2 & 2\alpha\beta \\ -2\alpha\beta & \beta^2-\alpha^2 \end{bmatrix} \begin{bmatrix} y_{12} \\ y_{11} \end{bmatrix} = \Psi^T \begin{bmatrix} y_{12} \\ y_{11} \end{bmatrix}\end{aligned}$$

and the first column of X by

$$\begin{aligned}\begin{bmatrix} x_{11} \\ -x_{12} \end{bmatrix} &= \begin{bmatrix} \beta^2-\alpha^2 & -2\alpha\beta \\ -2\alpha\beta & -\beta^2+\alpha^2 \end{bmatrix} \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \\ &= \begin{bmatrix} \beta^2-\alpha^2 & 2\alpha\beta \\ -2\alpha\beta & \beta^2-\alpha^2 \end{bmatrix} \begin{bmatrix} y_{11} \\ -y_{12} \end{bmatrix} = \Psi^T \begin{bmatrix} y_{11} \\ -y_{12} \end{bmatrix}.\end{aligned}$$

Hence $X = \Psi^T Y$. By taking the transpose of $X = \Psi^T Y$ and using the fact that $\Psi^T = \Psi^{-1}$, it follows that $Y = \Psi X$. Furthermore,

$$\begin{aligned}X\gamma + \gamma^T Y &= \Psi^T Y\gamma + \gamma^T Y \\ &= -(\gamma^T)^2 Y\gamma + \gamma^T Y = -\gamma^T(\gamma^T Y\gamma - Y) \\ &= 0,\end{aligned}$$

since $\gamma^T Y\gamma - Y = 0$ from Lemma 3.2.6. Hence from (3.10), $\gamma^T Z\gamma - Z = 0$. It now follows from Lemma 3.2.6 that $Z \in \mathcal{E}$. From Lemma 3.2.5 it follows that if Z is symmetric then $Z = cI$ for some constant c. □

Remarks

(i) $-\gamma^2 = \begin{bmatrix} \beta^2-\alpha^2 & -2\alpha\beta \\ 2\alpha\beta & \beta^2-\alpha^2 \end{bmatrix} = \Psi$. Hence, $\Psi^T = -(\gamma^T)^2$

(ii) $\Psi^{-1} = \Psi^T = \begin{bmatrix} \beta^2-\alpha^2 & 2\alpha\beta \\ -2\alpha\beta & \beta^2-\alpha^2 \end{bmatrix}$.

(iii) Since $(\Psi^{-1})^T = \Psi$, and $\det \Psi^n = \det\left(\begin{bmatrix} \beta^2-\alpha^2 & -2\alpha\beta \\ 2\alpha\beta & \beta^2-\alpha^2 \end{bmatrix}^n\right) = 1$ for all n, it follows that $((\Psi^n)^{-1})^T = \Psi^n$.

Example 3.2.3 (continued) From Lemma 3.2.5 and Lemma 3.2.7 we have $H_{1,1} = 0$ and $H_{2,2} = cI$. From the fact that $H_{2,2}$ is symmetric and from Lemma 3.2.8 it follows that $H_{1,2} \in \mathcal{E}$. Hence,

$$H = \left[\begin{array}{cc|cc} 0 & 0 & a & -b \\ 0 & 0 & b & a \\ \hline a & b & c & 0 \\ -b & a & 0 & c \end{array}\right]. \tag{3.13}$$

The following lemma also plays an important role in the sequel:

Lemma 3.2.9 *Let* $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, *and let* X, Y *and* W *be of the form* $aI + bH_0$ *and let* Z *be any* 2×2 *real matrix such that*

$$W + \gamma^T Y + X\gamma + \gamma^T Z\gamma - Z = 0. \tag{3.14}$$

Put $\Psi = -\gamma^2$. *Then,*

$$X = \Psi^T Y - \gamma^T W \quad \textit{and} \quad \gamma^T Z\gamma - Z = 0, \tag{3.15}$$

hence $Z \in \mathcal{E}$.

Proof. Take the trace of equation (3.14), that is

$$\operatorname{tr}(W + \gamma^T Y + X\gamma + \gamma^T Z\gamma - Z) = 0.$$

From the proof of Lemma 3.2.7, we have already seen that $\operatorname{tr}(\gamma^T Z\gamma - Z) = 0$. Thus

$$\operatorname{tr}(W) + \operatorname{tr}(\gamma^T Y) + \operatorname{tr}(X\gamma) = 0.$$

Since $W = \begin{bmatrix} w_{11} & w_{12} \\ -w_{12} & w_{11} \end{bmatrix}, Y = \begin{bmatrix} y_{11} & y_{12} \\ -y_{12} & y_{11} \end{bmatrix}$ and $X = \begin{bmatrix} x_{11} & x_{12} \\ -x_{12} & x_{11} \end{bmatrix}$ it follows that

$$w_{11} + \alpha y_{11} + \beta y_{12} + \alpha x_{11} - \beta x_{12} = 0. \tag{3.16}$$

Pre-multiplying equation (3.14) by H_0 and taking the trace again we have

$$\operatorname{tr}(H_0 W + H_0\gamma^T Y + H_0 X\gamma + H_0\gamma^T Z\gamma - H_0 Z) = 0.$$

As before, $\operatorname{tr}(H_0\gamma^T Z\gamma - H_0 Z) = 0$, by Lemma 3.2.7. The same computation as in (3.12) leads to

$$-w_{12} + \beta y_{11} - \alpha y_{12} - \alpha x_{12} - \beta x_{11} = 0. \tag{3.17}$$

Rewriting equations (3.16) and (3.17) in matrix form gives:

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} x_{11} \\ -x_{12} \end{bmatrix} = \begin{bmatrix} -\alpha & \beta \\ -\beta & -\alpha \end{bmatrix} \begin{bmatrix} y_{11} \\ -y_{12} \end{bmatrix} - \begin{bmatrix} w_{11} \\ -w_{12} \end{bmatrix}, \quad \text{i.e.}$$

$$\gamma \begin{bmatrix} x_{11} \\ -x_{12} \end{bmatrix} = -\gamma^T \begin{bmatrix} y_{11} \\ -y_{12} \end{bmatrix} - \begin{bmatrix} w_{11} \\ -w_{12} \end{bmatrix}$$

and

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{11} \end{bmatrix} = \begin{bmatrix} -\alpha & \beta \\ -\beta & -\alpha \end{bmatrix} \begin{bmatrix} y_{12} \\ y_{11} \end{bmatrix} - \begin{bmatrix} w_{12} \\ w_{11} \end{bmatrix}, \quad \text{i.e.}$$

$$\gamma \begin{bmatrix} x_{12} \\ x_{11} \end{bmatrix} = -\gamma^T \begin{bmatrix} y_{12} \\ y_{11} \end{bmatrix} - \begin{bmatrix} w_{12} \\ w_{11} \end{bmatrix}.$$

Hence,

$$\gamma X = -\gamma^T Y - W.$$

Pre-multiplying by $\gamma^{-1} = \gamma^T$, we have

$$\begin{aligned} X &= -(\gamma^T)^2 Y - \gamma^T W \\ &= \Psi^T Y - \gamma^T W. \end{aligned}$$

Substituting the above X into (3.14), we have

$$\begin{aligned} W + (\Psi^T Y - \gamma^T W)\gamma + \gamma^T Y + \gamma^T Z\gamma - Z &= 0, \quad \text{i.e.} \\ (W - \gamma^T W\gamma) + (\Psi^T Y\gamma + \gamma^T Y) + \gamma^T Z\gamma - Z &= 0, \quad \text{i.e.} \\ -(\gamma^T W\gamma - W) - \gamma^T(\gamma^T Y\gamma - Y) + \gamma^T Z\gamma - Z &= 0. \end{aligned}$$

By Lemma 3.2.6, we have that $\gamma^T Y\gamma - Y = 0$ and $\gamma^T W\gamma - W = 0$, whence $\gamma^T Z\gamma - Z = 0$. Again by Lemma 3.2.6, $Z \in \mathcal{E}$. □

3.3 Structure of H-expansive matrices

In this section we use examples to illustrate an algorithm to find a simple form for the matrix pair (A, H). We find the zero structure for the real matrix H, which in return gives us information needed to construct the invariant subspaces that we are ultimately looking for.

3.3.1 One Jordan block: one pair of complex eigenvalues

First, consider the case where A is a single real Jordan block with a single pair of eigenvalues $\alpha \pm i\beta$ on the unit circle. We assume that A is of (scalar) size $2n \times 2n$, that is, considered as a block matrix with 2×2 blocks, A is a $n \times n$ block matrix. Therefore, let

$$A = \begin{bmatrix} \gamma & I & & \\ & \gamma & \ddots & \\ & & \ddots & I \\ & & & \gamma \end{bmatrix},$$

where $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ and I is the 2×2 identity matrix. Also let $H = H^T$ be a real symmetric matrix. Remember that we shall write $H = [H_{i,j}]_{i,j=1}^n$, where each $H_{i,j}$ is a 2×2 block. If A is H-expansive, i.e. $A^T HA - H \geq 0$, we formulate the following result to obtain the simple form for the pair (A, H). Recall that $\Psi = -\gamma^2$.

Proposition 3.3.1 *Let A be a single real Jordan block with single pair of eigenvalues $\lambda = \alpha \pm i\beta$, each of multiplicity n, and $\alpha^2 + \beta^2 = 1$, $\beta \neq 0$. Let H be an invertible real symmetric $n \times n$ block matrix. Suppose A is H-expansive with respect to the indefinite inner product. Then, when n is odd, H has the following form:*

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & \dots & & \dots & 0 \quad c(\Psi^T)^{\frac{n-1}{2}} \\ 0 & & & & \iddots & \\ \vdots & & & & c(\Psi^T)^2 & \\ & & & c\Psi^T & & \\ & & cI & & & \\ \hline & & c\Psi & & & \\ \vdots & & c\Psi^2 & & & \\ 0 & \iddots & & & * & \\ c\Psi^{\frac{n-1}{2}} & & & & & \end{array}\right]. \tag{3.18}$$

Here c denotes the real value $(H_{\frac{n+1}{2},\frac{n+1}{2}})_{11}$. For even n we have

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & \dots & & \dots & 0 \quad (\Psi^{\frac{n-2}{2}})^T\mathcal{H} \\ 0 & & & & \iddots & \\ \vdots & & & & (\Psi^2)^T\mathcal{H} & \\ & & & \Psi^T\mathcal{H} & & \\ & & & \mathcal{H} & & \\ \hline & & \mathcal{H}^T & & & \\ \vdots & & \mathcal{H}^T\Psi & & & \\ & \mathcal{H}^T\Psi^2 & & & * & \\ 0 & \iddots & & & & \\ \mathcal{H}^T\Psi^{\frac{n-2}{2}} & & & & & \end{array}\right], \tag{3.19}$$

where $\mathcal{H} = H_{\frac{n}{2},\frac{n}{2}+1}$. The vertical and horizontal lines in the matrix divide the matrix into four blocks. The left upper block has size $m \times m$, with $m = [\frac{n+1}{2}]$, and consists of zero entries in the even case while in the odd case it has the entry cI in the right lower corner and all other entries are zero. In the lower right block are entries which are not of interest to us in the construction of the invariant subspaces. However, all other entries below the skew diagonal in the right upper block and the left lower block are entries which are in $\mathcal{E}$.

If n is even, then it follows for the block matrix $\mathcal{H} \in \mathcal{E}$, i.e., $\mathcal{H} = aI + bH_0$ where $a, b \in \mathbb{R}$ such that $a\alpha - b\beta \geq 0$.

Proof. The result follows from a repeated application of Lemma 3.2.5, Remark 3.2.1, Lemma 3.2.6, Lemma 3.2.7 and Lemma 3.2.8. The way these lemmas are applied is demonstrated in the construction of the simple forms for the cases $n = 5$ and $n = 6$ below. These cases can easily be extended to the general case.

Consider the case where n is odd first. The $(1,1)$-block entry of the block matrix $A^THA - H$ is given by:

$$(A^THA - H)_{1,1} = \gamma^T H_{1,1}\gamma - H_{1,1} \geq 0.$$

By Lemma 3.2.5, $\gamma^T H_{1,1}\gamma - H_{1,1} = 0$ and there exists a $c_1 \in \mathbb{R}$ such that $H_{1,1} = c_1 I$. Applying Remark 3.2.1, it follows that for all $j \geq 1$, $(\gamma^T H\gamma - H)_{1,j} = (\gamma^T H\gamma - H)_{j,1} = 0$.

For $j > 1$, knowing that all the entries in the first block row of $A^T HA - H$ are zero, we have

$$\begin{aligned} 0 &= (A^T HA - H)_{1,j} \\ &= \begin{bmatrix} \gamma^T & 0 & \dots & 0 & 0 \end{bmatrix} H \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ \gamma \\ 0 \\ \vdots \\ 0 \end{bmatrix} - H_{1,j} = \gamma^T H_{1,j-1} + \gamma^T H_{1,j}\gamma - H_{1,j}. \quad (3.20) \end{aligned}$$

Taking $j = 2$ we have from (3.20) that

$$\gamma^T H_{1,1} + \gamma^T H_{1,2}\gamma - H_{1,2} = 0.$$

Since $H_{1,1} = c_1 I$ we have by Lemma 3.2.7 that $H_{1,1} = 0$. Then by Lemma 3.2.6 we have $H_{1,2} \in \mathcal{E}$. Proceeding in this way for $i = 1$ and $j > 1$ we find $H_{1,j} = 0$ for $j = 1, \dots, n-1$ and $H_{1,n} \in \mathcal{E}$.

In general, for $i > 1$ and $j > 1$ we have

$$(A^T HA - H)_{i,j} = \begin{bmatrix} 0 & \dots & 0 & I & \gamma^T & 0 & \dots & 0 \end{bmatrix} H \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ \gamma \\ 0 \\ \vdots \\ 0 \end{bmatrix} - H_{i,j},$$

where I and γ^T in the first block row is in the $(i, i-1)$ and (i, i) position respectively and in the block column, I is in the $(j-1, j)$ position and γ in the (j, j) position. Hence,

$$(A^T HA - H)_{i,j} = H_{i-1,j-1} + H_{i-1,j}\gamma + \gamma^T H_{i,j-1} + \gamma^T H_{i,j}\gamma - H_{i,j}. \quad (3.21)$$

Consider the second block row in $A^T HA - A$, that is, let $i = 2$. Then (3.21) becomes

$$(A^T HA - H)_{2,j} = H_{1,j-1} + H_{1,j}\gamma + \gamma^T H_{2,j-1} + \gamma^T H_{2,j}\gamma - H_{2,j}. \quad (3.22)$$

For $j = 2$ in (3.22), we have

$$(A^T HA - H)_{2,2} = H_{1,1} + H_{1,2}\gamma + \gamma^T H_{2,1} + \gamma^T H_{2,2}\gamma - H_{2,2} \geq 0.$$

Since $H_{1,1} = 0$, $H_{1,2} = 0$ (from above) and $H_{2,1} = 0$ by symmetry, we have

$$(A^T HA - H)_{2,2} = \gamma^T H_{2,2}\gamma - H_{2,2} \geq 0.$$

It now follows from Lemma 3.2.5 that $\gamma^T H_{2,2}\gamma - H_{2,2} = 0$ and $H_{2,2} = c_2 I$ for some $c_2 \in \mathbb{R}$, i.e., $(A^T HA - H)_{2,2} = 0$. By Remark 3.2.1 we therefore have $(A^T HA - H)_{2,j} = 0$ for all $j = 1, 2, \ldots, n$.
Taking $j = 3$ in (3.22), it therefore follows that

$$H_{1,2} + H_{1,3}\gamma + \gamma^T H_{2,2} + \gamma^T H_{2,3}\gamma - H_{2,3} = 0.$$

Substituting $H_{1,2} = 0$ and $H_{1,3} = 0$ we have $\gamma^T H_{2,2} + \gamma^T H_{2,3}\gamma - H_{2,3} = 0$. Also, since $H_{2,2} = c_2 I$, we realize from Lemma 3.2.7 that $H_{2,2} = 0$ and therefore that

$$\gamma^T H_{2,3}\gamma - H_{2,3} = 0.$$

Now Lemma 3.2.6 kicks in to let us realize that $H_{2,3} \in \mathcal{E}$. Repeating the above procedure for $j = 4$ in (3.22), we soon realize that $H_{2,3} = 0$ and $H_{2,4} \in \mathcal{E}$.

We may continue in this manner, but for illustration of this procedure, we restrict ourselves to the case where $n = 5$. All arguments carry over to the general case. So, if $j = 5$ in (3.22), we have

$$H_{1,4} + H_{1,5}\gamma + \gamma^T H_{2,4} + \gamma^T H_{2,5}\gamma - H_{2,5} = 0.$$

Since $H_{1,4} = 0$, and $H_{1,5}, H_{2,4} \in \mathcal{E}$, it follows from Lemma 3.2.8 that $H_{1,5} = \Psi^T H_{2,4}$ and $H_{2,5} \in \mathcal{E}$.

For the next block row, let $i = 3$. For $j > 1$ in (3.21) we have

$$(A^T HA - H)_{3,j} = H_{2,j-1} + H_{2,j}\gamma + \gamma^T H_{3,j-1} + \gamma^T H_{3,j}\gamma - H_{3,j}. \tag{3.23}$$

Since $H = H^T$ we see that $H_{31} = H_{13} = 0$ and $H_{32} = H_{23} = 0$. For the other block entries in the third row we follow the same arguments as above and see that $(A^T HA - H)_{3,3} = \gamma^T H_{3,3}\gamma - H_{3,3} = 0$. So, $(A^T HA - H)_{3,j} = (A^T HA - H)_{j,3} = 0$ for all j. Applying Lemma 3.2.5 for $j = 3$ in (3.23) we find that $H_{3,3} = c_3 I$.
For $j = 4$ in (3.23) we have

$$H_{2,4}\gamma + \gamma^T H_{3,3} + \gamma^T H_{3,4}\gamma - H_{3,4} = 0.$$

Lemma 3.2.8 yields $H_{2,4} = \Psi^T H_{3,3} = c_3 \Psi^T$ and $H_{3,4} \in \mathcal{E}$. Hence,

$$H_{1,5} = \Psi^T H_{2,4} = (\Psi^T)^2 H_{3,3} = c_3 (\Psi^T)^2.$$

When $j = 5$ in (3.23) we obtain

$$H_{2,4} + H_{2,5}\gamma + \gamma^T H_{3,4} + \gamma^T H_{3,5}\gamma - H_{3,5} = 0.$$

From Lemma 3.2.9 it follows that $H_{3,5} \in \mathcal{E}$.

The algorithm stops for $(A^T HA - H)_{i,j}$ when $i, j > [\frac{n+1}{2}]$. We denote the remaining entries of the matrix by $*$. Each $*$ in the matrix below denotes a real 2×2 block. Also the zeros in the matrix below represent 2×2 blocks. Since $H = H^T$, we can find the other entries by symmetry. Thus,

$$H = \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 0 & c_3(\Psi^T)^2 \\ 0 & 0 & 0 & c_3\Psi^T & e_1 \\ 0 & 0 & c_3 I & e_2 & e_3 \\ \hline 0 & c_3\Psi & e_2^T & * & * \\ c_3\Psi^2 & e_1^T & e_3^T & * & * \end{array}\right],$$

where e_i denotes some 2×2 block entries which are in $\mathcal{E}$. The proof for the general case when n is odd, is completely analogous.

We now focus on the case when n is even. Let

$$A = \begin{bmatrix} \gamma & I & & \\ & \gamma & \ddots & \\ & & \ddots & I \\ & & & \gamma \end{bmatrix},$$

where $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$. Also, let $H = [H_{i,j}]_{i,j=1}^n$ and again remember that $H = H^T$. If A is H-expansive, we have

$$(A^T H A - H)_{1,1} = \gamma^T H_{1,1}\gamma - H_{1,1} \geq 0.$$

As before, from Lemma 3.2.5 it follows that $\gamma^T H_{1,1}\gamma - H_{1,1} = 0$ and $H_{1,1} = c_1 I$. Hence from Remark 3.2.1 we have $(A^T H A - H)_{1,j} = (A^T H A - H)_{j,1} = 0$ for all j. Like in the odd case (see (3.20)) we obtain:

$$\gamma^T H_{1,j-1} + \gamma^T H_{1,j}\gamma - H_{1,j} = 0. \tag{3.24}$$

Using (3.24), Lemma 3.2.6 and Lemma 3.2.7 as in the odd case, we have

$$H_{1,j} = 0 \quad \text{for} \quad j = 1, 2, \ldots, n-1 \quad \text{and}$$

$$H_{1,n} \in \mathcal{E}.$$

In general, for $i > 1$ and $j > 1$ we again have

$$(A^T H A - H)_{i,j} = H_{i-1,j-1} + H_{i-1,j}\gamma + \gamma^T H_{i,j-1} + \gamma^T H_{i,j}\gamma - H_{i,j}. \tag{3.25}$$

For $i = 2$ and $j > 1$, equation (3.25) becomes

$$(A^T H A - H)_{2,j} = H_{1,j-1} + H_{1,j}\gamma + \gamma^T H_{2,j-1} + \gamma^T H_{2,j}\gamma - H_{2,j}. \tag{3.26}$$

Taking $j = 2, \ldots, n-1$ we have the same result as in the odd case, i.e., $(A^T H A - H)_{2,j} = 0$. As an illustration when we take $n = 6$, we have for $j = 5$ in (3.26)

$$H_{1,4} + H_{1,5}\gamma + \gamma^T H_{2,4} + \gamma^T H_{2,5}\gamma - H_{2,5} = 0.$$

Since $H_{1,4} = 0$, $H_{1,5} = 0$ and $H_{2,4} \in \mathcal{E}$, it follows from Lemma 3.2.7 that $H_{2,4} = 0$ and from Lemma 3.2.6 we obtain $H_{2,5} \in \mathcal{E}$. Similarly, when $j = 6$ in (3.26), it follows that

$$H_{1,5} + H_{1,6}\gamma + \gamma^T H_{2,5} + \gamma^T H_{2,6}\gamma - H_{2,6} = 0.$$

We know $H_{1,5} = 0$, $H_{1,6}$ and $H_{2,5}$ are in the set $\mathcal{E}$, hence it follows from Lemma 3.2.8 that

$$H_{1,6} = \Psi^T H_{2,5} \quad \text{and} \quad H_{2,6} \in \mathcal{E}.$$

We continue in the same manner for $i = 3$ and $j = 4, 5, 6$. Then, when $j = 4$ in (3.25) we have

$$(A^T H A - H)_{3,4} = \gamma^T H_{3,3} + \gamma^T H_{3,4}\gamma - H_{3,4} = 0.$$

From Lemma 3.2.7 we have $H_{3,3} = 0$ and from Lemma 3.2.6 we obtain $H_{3,4} \in \mathcal{E}$. If $j = 5$ in (3.25) we see from Lemma 3.2.8 that $H_{2,5} = \Psi^T H_{3,4}$ and $H_{3,5} \in \mathcal{E}$. Hence,

$$H_{1,6} = (\Psi^2)^T H_{3,4}.$$

If we let $j = 6$ in (3.25), then it follows from Lemma 3.2.9 that $H_{3,6} \in \mathcal{E}$.

The remaining cross-diagonal entries of H can be found by symmetry. If we let $H_{3,4} = \mathcal{H} \in \mathcal{E}$, then we have in the even case for $n = 6$ that

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & (\Psi^2)^T\mathcal{H} \\ 0 & 0 & 0 & 0 & \Psi^T\mathcal{H} & e_1 \\ 0 & 0 & 0 & \mathcal{H} & e_2 & e_3 \\ \hline 0 & 0 & \mathcal{H}^T & * & * & * \\ 0 & \mathcal{H}^T\Psi & e_2^T & * & * & * \\ \mathcal{H}^T\Psi^2 & e_1^T & e_3^T & * & * & * \end{array}\right],$$

where as in the odd case e_i denotes some 2×2 block entries which are in $\mathcal{E}$.

We will now prove the last part of Proposition 3.3.1 for even n. When we consider the $(4, 4)$ block entry of (3.25) we only know that this entry should be greater or equal to zero. But we obtain some useful information in doing so. Thus,

$$\begin{aligned} 0 \leq (A^T H A - H)_{4,4} &= H_{3,3} + H_{3,4}\gamma + \gamma^T H_{4,3} + \gamma^T H_{4,4}\gamma - H_{4,4} \\ &= H_{3,4}\gamma + \gamma^T H_{3,4}^T + \gamma^T H_{4,4}\gamma - H_{4,4} \end{aligned} \tag{3.27}$$

Since $H_{3,4} \in \mathcal{E}$, the first two terms reduces to the following expression in terms of α and β, that is

$$H_{3,4}\gamma + \gamma^T H_{3,4}^T = 2(a\alpha - b\beta)I,$$

for some $a, b \in \mathbb{R}$ and where we recall that $\lambda = \alpha \pm i\beta$.

If $a\alpha - b\beta < 0$, then clearly for (3.27) to hold, we should have

$$\gamma^T H_{4,4}\gamma - H_{4,4} \geq 0.$$

But then by Lemma 3.2.5 we have

$$\gamma^T H_{4,4}\gamma - H_{4,4} = 0.$$

This then contradicts the fact that $(A^T H A - H)_{4,4} \geq 0$ in (3.27). Therefore, $a\alpha - b\beta \geq 0$ should hold.

Applying the same reasoning as in the proof of Lemma 3.2.5 for the last two terms above, i.e., $\gamma^T H_{4,4}\gamma - H_{4,4}$ we see that (3.27) reduces to the following observation:

$$a\alpha - b\beta \geq 0.$$

□

In the remainder of this section, we consider some special cases with only one Jordan block.

Special case 1 ($\lambda = \pm i$)

We consider the special case when $\lambda = \pm i$. Now γ will be the 2×2 matrix

$$\gamma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = H_0.$$

As before, we have to consider the odd case (i.e. when n is odd) and the even case (when n is even) separately. Again we restrict ourselves to a $n \times n$ Jordan block matrix

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & & & \\ -1 & 0 & 0 & 1 & & & \\ & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 1 & 0 \\ & & & & & 0 & 1 \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{bmatrix}.$$

In this case $\Psi = -\gamma^2 = I$. Let $H = [H_{i,j}]_{i,j=1}^n$. If n is odd we immediately have

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & \dots & & 0 & cI \\ 0 & & & & \iddots & \\ \vdots & & & & cI & \\ & & & cI & & \\ & & cI & & & \\ \hline & & & & & \\ & cI & & & * & \\ 0 & \iddots & & & & \\ cI & & & & & \end{array}\right],$$

and when n is even, we have

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & \dots & & 0 & \mathcal{H} \\ 0 & & & & \iddots & \\ \vdots & & & \mathcal{H} & & \\ & & & \mathcal{H} & & \\ & & & \mathcal{H} & & \\ \hline & & \mathcal{H}^T & & & \\ & \mathcal{H}^T & & & & \\ \mathcal{H}^T & & & & * & \\ 0 & \iddots & & & & \\ \mathcal{H}^T & & & & & \end{array}\right],$$

where $\mathcal{H} \in \mathcal{E}$.

Special case 2 ($\lambda = \pm 1$)

It may happen that the unimodular eigenvalue of our matrix A is real, that is $\lambda = 1$ or $\lambda = -1$. The eigenvalues therefore do not appear in conjugate pairs, hence both A and H are no longer $n \times n$ block matrices. When $\lambda = 1$, we let A be a $n \times n$ matrix of the form

$$A = \begin{bmatrix} 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & 1 & 1 \\ & & & & & 1 \end{bmatrix}.$$

Let $H = [h_{ij}]_{i,j=1}^n$ and $H = H^T$. When A is H-expansive, $(A^THA - H)_{11} = h_{11} - h_{11} = 0$. Hence, $(A^THA - H)_{1j} = (A^THA - H)_{j1} = 0$ for all j. It is therefore clear that $h_{1j} = 0$ for $j = 1, \dots, n-1$.

For $i > 1$ and $j > 1$ we have

$$0 = (A^THA - H)_{ij} = h_{i-1,j-1} + h_{i-1,j} + h_{i,j-1}. \tag{3.28}$$

By direct calculation we have for $i = 2$ that

$$h_{2,j} = 0 \quad \text{for} \quad j = 1, \dots, n-2 \quad \text{and} \quad h_{2,n-1} = -h_{1,n}.$$

In general we obtain

$$h_{k,n-k+1} = (-1)^{k-1} h_{1,n} \quad \text{for} \quad k = 1, 2, \dots, n.$$

We can continue in this way as long as $i \leq [\frac{n+1}{2}]$, because when $i > [\frac{n+1}{2}]$ we only know that $(A^THA - H)_{ii} \geq 0$. Hence

$$(A^THA - H)_{ij} = 0 \quad \text{for} \quad 1 \leq i \leq [\frac{n+1}{2}] \quad \text{or} \quad 1 \leq j \leq [\frac{n+1}{2}].$$

Therefore, the corresponding matrix H has the following form for odd n:

$$H = \left[\begin{array}{cccc|ccc} 0 & 0 & \dots & & & \dots & 0 & c \\ 0 & & & & & & \iddots & \\ \vdots & & & & & & & \\ & & & & & (-1)^{\frac{n-3}{2}}c & & \\ & & & (-1)^{\frac{n-1}{2}}c & & & & \\ \hline & & (-1)^{\frac{n+1}{2}}c & & & & & \\ \vdots & & & & & & * & \\ 0 & \iddots & & & & & & \\ (-1)^{n-1}c & & & & & & & \end{array}\right]. \tag{3.29}$$

When n is even, we obtain the following formula

$$h_{k,n-k+1} = (-1)^{k-1}h_{1,n} \quad \text{for} \quad k = 1, 2, \dots, \frac{n}{2}.$$

Thus, in the case when n is even, the following simplified form for the matrix H is obtained:

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & \dots & & \dots & 0 & c \\ 0 & & & & & \iddots & \\ \vdots & & & & & & \\ & & & & (-1)^{\frac{n-4}{2}}c & & \\ & & & (-1)^{\frac{n-2}{2}}c & & & \\ \hline & & (-1)^{\frac{n-2}{2}}c & & & & \\ \vdots & (-1)^{\frac{n-4}{2}}c & & & & & \\ & & & & * & & \\ 0 & \iddots & & & & & \\ c & & & & & & \end{array}\right]. \tag{3.30}$$

In both cases $c = h_{1,n}$ and $k = 1, \dots, n$. Take note that $(-1)^{\frac{n-2}{2}}c > 0$. To see this, take $i = j = \frac{n}{2} + 1$ in (3.28).

When $\lambda = -1$ the matrix A has the following Jordan form

$$A = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \\ & & & & & -1 \end{bmatrix}.$$

The only difference is in (3.28), which becomes

$$(A^T H A - H)_{ij} = h_{i-1,j-1} - h_{i-1,j} - h_{i,j-1}.$$

When we apply the corresponding lemmas to determine the entries of the matrix H we find that they are exactly the same as in the case when $\lambda = 1$. So, the form of the matrix H is as in (3.29) and (3.30) for odd and even n, respectively. But now for even n we have $(-1)^{\frac{n-2}{2}}c < 0$. As before this is seen by taking $i = j = \frac{n}{2} + 1$ in (3.28).

Note that in the above special cases for the specific chosen A and H it will always hold that $A^T HA - H \neq 0$.

3.3.2 More Jordan blocks: one pair of complex eigenvalues

In this section we discuss a more general case, where we consider the case that the matrix A consists of more than one real Jordan block, but still with only one pair of eigenvalues, $\lambda = \alpha \pm i\beta$, $\beta \neq 0$, on the unit circle. To set up notation, let $A = J_1 \oplus \ldots \oplus J_l$, where J_j is a real Jordan block with eigenvalues $\alpha \pm i\beta$, each of multiplicity n_j. It is not excluded that $n_j = n_k$ for $j \neq k$. Hence, A may consist of Jordan blocks which can be all of even size, or all of odd size, or of both even and odd sizes. We consider the different cases. Once again we explain the construction of the simple forms for the matrix H by means of examples.

Firstly, let us consider a matrix A consisting of three real Jordan blocks. Let

$$A = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{bmatrix}$$

be in real Jordan form. Let each J_p, for $p = 1, 2, 3$, be of the following form

$$J_p = \begin{bmatrix} \gamma & I & & & \\ & \gamma & I & & \\ & & \ddots & \ddots & \\ & & & \gamma & I \\ & & & & \gamma \end{bmatrix},$$

with $\gamma = \alpha I + \beta H_0$ as before. Furthermore, let J_1 be a $n_1 \times n_1$, J_2 a $n_2 \times n_2$ and J_3 a $n_3 \times n_3$ Jordan block matrix, so as scalar matrix J_p is of size $2n_p \times 2n_p$. Put $N = n_1 + n_2 + n_3$. As before the block matrix H is a real symmetric block matrix of size $N \times N$. Let $H = [H_{i,j}]_{i,j=1}^3$, where each $H_{i,j}$ is of appropriate size and consists of 2×2 block entries. If A is H-expansive, we have

$$\begin{aligned} 0 &\leq A^T HA - H \\ &= \begin{bmatrix} J_1^T H_{1,1} J_1 - H_{1,1} & J_1^T H_{1,2} J_2 - H_{1,2} & J_1^T H_{1,3} J_3 - H_{1,3} \\ J_2^T H_{2,1} J_1 - H_{2,1} & J_2^T H_{2,2} J_2 - H_{2,2} & J_2^T H_{2,3} J_3 - H_{2,3} \\ J_3^T H_{3,1} J_1 - H_{3,1} & J_3^T H_{3,2} J_2 - H_{3,2} & J_3^T H_{3,3} J_3 - H_{3,3} \end{bmatrix}. \end{aligned}$$

Hence, for A to be H-expansive, it must hold that $J_p^T H_{p,p} J_p - H_{p,p} \geq 0$ for each $p = 1, 2, 3$. From the previous section a form for the matrix $H_{p,p}$ is already known for both odd and even n_p. Therefore we only need to consider the blocks $J_p^T H_{p,q} J_q - H_{p,q}$ when $p \neq q$.

Since $(J_1^T H_{1,1} J_1 - H_{1,1})_{i,j} = 0$ for $1 \leq i \leq [\frac{n_1+1}{2}]$ or $1 \leq j \leq [\frac{n_1+1}{2}]$, we have that

$$
\begin{aligned}
&(A^T H A - H)_{i,j} = 0, \\
\text{for} \quad & i = 1, \ldots, \left[\frac{n_1+1}{2}\right] \quad \text{and} \quad j = 1, \ldots, N \\
\text{or} \quad & j = 1, \ldots, \left[\frac{n_1+1}{2}\right] \quad \text{and} \quad i = 1, \ldots, N.
\end{aligned}
$$

Also, because $(J_2^T H_{2,2} J_2 - H_{2,2})_{i,j} = 0$ for $n_1+1 \le i \le [\frac{n_1+n_2+1}{2}]$ or $n_1+1 \le j \le [\frac{n_1+n_2+1}{2}]$, we have

$$
\begin{aligned}
&(A^T H A - H)_{i,j} = 0, \\
\text{for} \quad & i = n_1+1, \ldots, n_1 + \left[\frac{n_2+1}{2}\right] \quad \text{and} \quad j = 1, \ldots, N \\
\text{or} \quad & j = n_1+1, \ldots, n_1 + \left[\frac{n_2+1}{2}\right] \quad \text{and} \quad i = 1, \ldots, N.
\end{aligned}
$$

Similarly, since $(J_3^T H_{3,3} J_3 - H_{3,3})_{i,j} = 0$ for $n_1+n_2+1 \le i \le [\frac{n_1+n_2+n_3+1}{2}]$ or $n_1+n_2+1 \le j \le [\frac{n_1+n_2+n_3+1}{2}]$, we have

$$
\begin{aligned}
&(A^T H A - H)_{i,j} = 0, \\
\text{for} \quad & i = n_1+n_2+1, \ldots, n_1+n_2 + \left[\frac{n_3+1}{2}\right] \quad \text{and} \quad j = 1, \ldots, N \\
\text{or} \quad & j = n_1+n_2+1, \ldots, n_1+n_2 + \left[\frac{n_3+1}{2}\right] \quad \text{and} \quad i = 1, \ldots, N.
\end{aligned}
$$

If in our example we let $n_1 = 5, n_2 = 4$ and $n_3 = 3$, then the block matrix $A^T H A - H$ has 2×2 block zero entries in the following positions:

$$
A^T H A - H = \left[\begin{array}{ccccc|cccc|ccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & & & 0 & 0 & & & 0 & 0 & \\
0 & 0 & 0 & & & 0 & 0 & & & 0 & 0 & \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & & & 0 & 0 & & & 0 & 0 & \\
0 & 0 & 0 & & & 0 & 0 & & & 0 & 0 & \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & & & 0 & 0 & & & 0 & 0 &
\end{array}\right].
$$

The blocks with blanks on the main diagonal must be positive semi-definite. The other blocks with blanks are not specified. The form of the blocks on the main diagonal of H, for even and odd n are known. Since $H = H^T$, we now only have to consider the blocks

below the main diagonal which are not necessarily square matrices. We use the same algorithm as in the previous construction.

Firstly, let us consider the block $J_2^T H_{2,1} J_1 - H_{2,1}$. Let $H_{2,1} = [\tilde{H}_{i,j}]$ for $i = 1, \ldots, n_2$ and $j = 1, \ldots, n_1$. It follows from the above discussion that $\gamma^T \tilde{H}_{1,1} \gamma - \tilde{H}_{1,1} = 0$. From Lemma 3.2.6 we know that $\tilde{H}_{1,1} \in \mathcal{E}$. Compare to the first few paragraphs immediately after the proof of Proposition 3.3.1. For $i = 1$ and $j > 1$ we have

$$0 = \gamma^T \tilde{H}_{1,j-1} + \gamma^T \tilde{H}_{1,j} \gamma - \tilde{H}_{1,j}.$$

We apply Lemma 3.2.6 and Lemma 3.2.7. As expected, we have $\tilde{H}_{1,j} = 0$ for $j = 1, \ldots, n_1 - 1$ and $\tilde{H}_{1,n_1} \in \mathcal{E}$. From the fact that $(J_2^T H_{2,1} J_1 - H_{2,1})_{2,1} = 0$, it follows that $\tilde{H}_{1,1}\gamma + \gamma^T \tilde{H}_{2,1} \gamma - \tilde{H}_{2,1} = 0$. And since $\tilde{H}_{1,1} = 0$, it follows from Lemma 3.2.6 that $\tilde{H}_{2,1} \in \mathcal{E}$.

If $i > 1$ and $j > 1$ we have

$$(J_2^T H_{2,1} J_1 - H_{2,1})_{i,j} = \tilde{H}_{i-1,j-1} + \tilde{H}_{i-1,j}\gamma + \gamma^T \tilde{H}_{i,j-1} + \gamma^T \tilde{H}_{i,j} \gamma - \tilde{H}_{i,j}. \tag{3.31}$$

For $i = 2$ and $j = 2, \ldots, n_1 - 1$ iteratively in (3.31) together with Lemma 3.2.7, we have that $\tilde{H}_{2,1} = \tilde{H}_{2,2} = \tilde{H}_{2,3} = 0$ and $\tilde{H}_{2,4} \in \mathcal{E}$. When $j = n_1 = 5$ in (3.31) we have from Lemma 3.2.8 that

$$\tilde{H}_{1,5} = \Psi^T \tilde{H}_{2,4}.$$

Increasing i by one and successively increasing j by one we have as before that: $\tilde{H}_{3,1} = \tilde{H}_{3,2} = 0$ and $\tilde{H}_{3,3} \in \mathcal{E}$. We stop here, because we do not know if the entry $(J_2^T H_{2,1} J_1 - H_{2,1})_{3,4}$ is zero.

When $i = 4$ and $1 \leq j \leq n_1 - 2$ we have that $\tilde{H}_{4,1} = 0$, $\tilde{H}_{4,2} \in \mathcal{E}$ and

$$\tilde{H}_{3,3} = \Psi^T \tilde{H}_{4,2}.$$

Again we stop, because it is not clear whether or not the next entry in the block is zero.

Hence, we have a structure for the block $H_{2,1}$. We now proceed with the remaining blocks under the main diagonal. Let us consider the block $J_3^T H_{3,1} J_1 - H_{3,1}$. In what follows we denote the entries of $H_{3,1}$ by $H'_{i,j}$. Note that $H_{3,1}$ in our example is a $n_3 \times n_1$ matrix. For this block we let $i = 1, \ldots, n_3$ and $j = 1, \ldots, n_1$. We know that

$$\begin{gathered} (J_3^T H_{3,1} J_1 - H_{3,1})_{i,j} = 0 \\ \text{for} \quad i = 1, \ldots, \left[\frac{n_3 + 1}{2}\right] \quad \text{and} \quad j = 1, \ldots, n_1 \\ \text{or} \quad j = 1, \ldots, \left[\frac{n_1 + 1}{2}\right] \quad \text{and} \quad i = 1, \ldots, n_3. \end{gathered}$$

Following the same algorithm as for the previous block, we find that $H'_{1,j} = 0$ for $j = 1, \ldots, n_1 - 1$ and $H'_{1,n_1} \in \mathcal{E}$. We use equation (3.31) for this and all the remaining blocks. So, as before, we can increase i and run through all possible j's before we increase i again. Thus we have $H'_{2,j} = 0$ for $j = 1, \ldots, n_1 - 2$ and $H'_{2,4} \in \mathcal{E}$ such that

$$H'_{1,5} = \Psi^T H'_{2,4}.$$

We also have $H'_{3,j} = 0$ for $j = 1, \ldots, n_1 - 3$ and $H'_{3,3} \in \mathcal{E}$. Hence we have a structure for the block $H_{3,1}$.

It remains to consider the block $J_3^T H_{3,2} J_2 - H_{3,2}$. We denote the block entries of $H_{3,2}$ by $\tilde{H}'_{i,j}$. Here we have that

$$\begin{aligned} &(J_3^T H_{3,2} J_2 - H_{3,2})_{i,j} = 0, \\ \text{for} \quad & i = 1, \ldots, \left[\frac{n_3+1}{2}\right] \quad \text{and} \quad j = 1, \ldots, n_2 \\ \text{or} \quad & j = 1, \ldots, \left[\frac{n_2+1}{2}\right] \quad \text{and} \quad i = 1, \ldots, n_3. \end{aligned}$$

Applying Lemma 3.2.6, Lemma 3.2.7 and Lemma 3.2.8 together with equation (3.31), we have: $\tilde{H}'_{1,j} = 0$ for $j = 1, \ldots, n_2 - 1$, $\tilde{H}'_{1,n_2} \in \mathcal{E}$. Applying this again for $i = 2$ we obtain $\tilde{H}'_{2,j} = 0$ for $j = 1, \ldots, n_2 - 2$, $\tilde{H}'_{2,n_2-1} \in \mathcal{E}$ and $\tilde{H}'_{1,n_2} = \Psi^T \tilde{H}'_{2,n_2-1}$. For $i = 3$ we get $\tilde{H}'_{3,1} = 0$ and $\tilde{H}'_{3,2} \in \mathcal{E}$.

Since we have $H = H^T$, we immediately find the entries for H in the blocks above the main diagonal blocks. Note that the (1,1)-block and the (3,3)-block are derived from (3.18), while the (2,2)-block is obtained from (3.19). Although we have obtained the (2,1), (3,1) and the (3,2)-blocks, we only exhibit the (2,1)-block. For later application we are only interested in the position of the zeros within all the blocks. Hence we get the following form for H:

$$\left[\begin{array}{cccccc|ccccc|ccc} 0 & 0 & 0 & 0 & c(\Psi^2)^T & & 0 & 0 & 0 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 0 & c\Psi^T & * & & 0 & 0 & 0 & \tilde{H}_{4,2}^T & & 0 & 0 & 0 \\ 0 & 0 & cI & * & * & & 0 & 0 & \tilde{H}_{4,2}^T\Psi & * & & 0 & 0 & * \\ 0 & c\Psi & * & * & * & & 0 & \tilde{H}_{2,4}^T & * & * & & 0 & * & * \\ c\Psi^2 & * & * & * & * & & \tilde{H}_{2,4}^T\Psi & * & * & * & & * & * & * \\ \hline 0 & 0 & 0 & 0 & \Psi^T\tilde{H}_{2,4} & & 0 & 0 & 0 & \Psi^T\overline{H}_{2,3} & & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{H}_{2,4} & * & & 0 & 0 & \overline{H}_{2,3} & * & & 0 & 0 & * \\ 0 & 0 & \Psi^T\tilde{H}_{4,2} & * & * & & 0 & \overline{H}_{2,3}^T & * & * & & 0 & * & * \\ 0 & \tilde{H}_{4,2} & * & * & * & & \overline{H}_{2,3}^T\Psi & * & * & * & & * & * & * \\ \hline 0 & 0 & 0 & 0 & * & & 0 & 0 & 0 & * & & 0 & 0 & \tilde{c}\Psi^T \\ 0 & 0 & 0 & * & * & & 0 & 0 & * & * & & 0 & \tilde{c}I & * \\ 0 & 0 & * & * & * & & 0 & * & * & * & & \tilde{c}\Psi & * & * \end{array}\right]. \tag{3.32}$$

Finally we also consider the special case where $\lambda = \pm 1$. As we have seen in the section where this special case was considered with only one Jordan block, our matrix pair (A, H) consist of scalar entries. Hence, if A consist of more Jordan blocks, we find on the main diagonal blocks for H either the form as in (3.29) or (3.30), depending on n.

Below the main diagonal blocks of H, we find something similar to (3.32), but now with scalar entries on the skew diagonals. These entries alternate in sign and depend in each block on the $(1, n_j)$-entry. For example, in the (2,1)-block the entries on the main skew diagonal is given by

$$h_{k,n_1-k+1} = (-1)^{k-1} h_{1,n_1} \quad \text{for} \quad k = 1, \ldots n_2$$

where n_1 is the dimension of the first Jordan block and n_2 the dimension of the second Jordan block.

The case when $\lambda = 1$ or $\lambda = -1$ is more complicated. This is illustrated already by the canonical form for real H-unitary matrices, as given in [66], Section 7. The difficulties pertain to Jordan blocks with eigenvalues ± 1 and equal even size. For the purpose of constructing invariant maximal H-semidefinite subspaces this is not a problem, as we shall see later on. The precise structure of the blocks in H corresponding to Jordan blocks in A with eigenvalues ± 1 and even size is immaterial for this purpose.

However, from point of view of constructing a simple form it would be of interest to delve into this case a bit deeper. This is left as a future topic of investigation. Just to show how complicated matters can be, we close this section with an illustrative example. The example shows that a simple form must of necessity depend very much on the precise nature of the matrix $D = A^T H A - H$.

Let us analyze the case of two blocks of size two, with eigenvalue one, that is,

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & t & 0 & v \\ t & u & w & x \\ 0 & w & 0 & y \\ v & x & y & z \end{bmatrix}$$

for some real numbers t, u, v, w, x, y and z.

We introduce

$$D = A^T H A - H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\,t & 0 & w+v \\ 0 & 0 & 0 & 0 \\ 0 & w+v & 0 & 2\,y \end{bmatrix}.$$

Recall that by assumption $D \geq 0$.

The effect of a change of Jordan basis is to replace H by $H_1 = S^T H S$, where S is an invertible matrix of the following form:

$$S = \begin{bmatrix} a & b & c & d \\ 0 & a & 0 & c \\ e & f & g & h \\ 0 & e & 0 & g \end{bmatrix}.$$

By invertibility of S we have $ag - ce \neq 0$. Then $A_1 = S^{-1} A S = A$. Also, we shall compute $D_1 = S^T D S$. Note that this does not have an effect on the rank of D.

We have

$$H_1 = \begin{bmatrix} 0 & t_1 & 0 & v_1 \\ t_1 & u_1 & w_1 & x_1 \\ 0 & w_1 & 0 & y_1 \\ v_1 & x_1 & y_1 & z_1 \end{bmatrix},$$

where

$$
\begin{aligned}
t_1 &= (at+ew)a+(av+ey)e,\\
v_1 &= (at+ew)c+(av+ey)g,\\
w_1 &= (at+ev)c+(aw+ey)g,\\
y_1 &= (ct+gw)c+(cv+gy)g,\\
u_1 &= (at+ev)\,b+(bt+au+fw+ex)\,a+(aw+ey)\,f+(bv+ax+fy+ez)\,e,\\
x_1 &= (at+ev)\,d+(bt+au+fw+ex)\,c+(aw+ey)\,h+(bv+ax+fy+ez)\,g,\\
z_1 &= (ct+gv)\,d+(dt+cu+hw+gx)\,c+(cw+gy)\,h+(dv+cx+hy+gz)\,g.
\end{aligned}
$$

Also,

$$
D_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2t_1 & 0 & w_1+v_1 \\ 0 & 0 & 0 & 0 \\ 0 & w_1+v_1 & 0 & 2y_1 \end{bmatrix}
$$

Observe that

$$
\begin{bmatrix} 2t_1 & w_1+v_1 \\ w_1+v_1 & 2y_1 \end{bmatrix} = \begin{bmatrix} a & e \\ c & g \end{bmatrix} \begin{bmatrix} 2t & w+v \\ w+v & 2y \end{bmatrix} \begin{bmatrix} a & c \\ e & g \end{bmatrix}.
$$

Now by Sylvester's law of inertia we can choose a, c, e and g such that $ag - ce \neq 0$ and

$$
\begin{bmatrix} 2t_1 & w_1+v_1 \\ w_1+v_1 & 2y_1 \end{bmatrix} = \begin{cases} 0 & \text{if rank } D = 0, \\ \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} & \text{if rank } D = 1, \\ 2I & \text{if rank } D = 2. \end{cases}
$$

Suppose we have chosen a, c, e, g such that this holds. Then, in all cases we have $v_1 = w_1$, and further we have, in case rank $D = 0$, also $t_1 = 0, y_1 = 0$. In case rank $D = 1$ we have $t_1 = 1, y_1 = 0$, in case rank $D = 2$ we have $t_1 = y_1 = 1$. Thus, we have

$$
H_1 = \begin{cases} \begin{bmatrix} 0 & 0 & 0 & v_1 \\ 0 & u_1 & -v_1 & x_1 \\ 0 & -v_1 & 0 & 0 \\ v_1 & x_1 & 0 & z_1 \end{bmatrix} & \text{if rank } D = 0, \\ \begin{bmatrix} 0 & 1 & 0 & v_1 \\ 1 & u_1 & -v_1 & x_1 \\ 0 & -v_1 & 0 & 0 \\ v_1 & x_1 & 0 & z_1 \end{bmatrix} & \text{if rank } D = 1, \\ \begin{bmatrix} 0 & 1 & 0 & v_1 \\ 1 & u_1 & -v_1 & x_1 \\ 0 & -v_1 & 0 & 1 \\ v_1 & x_1 & 1 & z_1 \end{bmatrix} & \text{if rank } D = 2. \end{cases}
$$

Now let us rewrite the formulas for the remaining variables. We have $v_1 = (ct + gv)a + (cw + gy)e$, and

$$\begin{aligned} u_1 &= 2(at + ev)b + 2(aw + ey)f + (a^2u + e^2z + 2aex) \\ z_1 &= 2(ct + gv)d + 2(cw + gy)h + (c^2u + g^2z + 2gcx) \\ x_1 &= (at + ev)d + (cw + gy)f + (ct + gv)b + (aw + ey)h \\ &\quad + (au + ex)c + (ax + ez)g \end{aligned}$$

Rewriting this as a set of equations in b, d, f, h given a, c, e, g, we arrive at

$$\begin{bmatrix} u_1 \\ z_1 \\ x_1 \end{bmatrix} = X \begin{bmatrix} b \\ d \\ f \\ h \end{bmatrix} + \xi,$$

where

$$X = \begin{bmatrix} 2(at + ev) & 0 & 2(aw + ey) & 0 \\ 0 & 2(ct + gv) & 0 & 2(cw + gy) \\ ct + gv & at + gv & cw + gy & aw + ey \end{bmatrix},$$

and

$$\xi = \begin{bmatrix} a^2u + e^2z + 2aex \\ c^2u + g^2z + 2gcx \\ (au + ex)c + (ax + ez)g \end{bmatrix}.$$

We claim that X is onto. To see this, we need to show that X has rank three. We show that the first and the third column are independent. If not, then

$$(at + ev)(cw + gy) - (aw + ey)(ct + gv) = 0,$$

which, after some computation, is equivalent to

$$(ag - ec)(ty - vw) = 0.$$

Now both terms are non-zero: $ag - ec \neq 0$ since S is invertible, and $ty - vw \neq 0$ since H is invertible. So the first and third column are independent. Likewise, the second and fourth column are independent. The first column is non-zero. Indeed, if it were zero, then $at + ev = 0$ and $ct + gv = 0$, that is

$$\begin{bmatrix} a & e \\ c & g \end{bmatrix} \begin{bmatrix} t \\ v \end{bmatrix} = 0.$$

Since S is invertible $ag - ce \neq 0$, and so this would imply $t = v = 0$. But then H is not invertible. Thus the first column is non-zero. In a similar manner the other three columns are non-zero as well. Hence the rank of X is three.

It then follows that there exist b, d, f, h such that $u_1 = 0, z_1 = 0, x_1 = 0$. Thus we can reduce to the case where

$$H_1 = \begin{cases} \begin{bmatrix} 0 & 0 & 0 & v_1 \\ 0 & 0 & -v_1 & 0 \\ 0 & -v_1 & 0 & 0 \\ v_1 & 0 & 0 & 0 \end{bmatrix} & \text{if rank } D = 0, \\ \begin{bmatrix} 0 & 1 & 0 & v_1 \\ 1 & 0 & -v_1 & 0 \\ 0 & -v_1 & 0 & 0 \\ v_1 & 0 & 0 & 0 \end{bmatrix} & \text{if rank } D = 1, \\ \begin{bmatrix} 0 & 1 & 0 & v_1 \\ 1 & 0 & -v_1 & 0 \\ 0 & -v_1 & 0 & 1 \\ v_1 & 0 & 1 & 0 \end{bmatrix} & \text{if rank } D = 2. \end{cases}$$

Our next step is to do a second similarity, this time with $S_1 = \begin{bmatrix} a_1 I_2 & c_1 I_2 \\ e_1 I_2 & g_1 I_2 \end{bmatrix}$, and we put $H_2 = S_1^T H_1 S_1$ with H_1 in the form we just obtained. We assume S_1 is invertible. In case rank $D = 0$, we obtain

$$H_2 = \begin{bmatrix} 0 & 0 & 0 & v_2 \\ 0 & 0 & -v_2 & 0 \\ 0 & -v_2 & 0 & 0 \\ v_2 & 0 & 0 & 0 \end{bmatrix}$$

where $v_2 = (a_1 g_1 - c_1 e_1) v_1$. Clearly, by choosing $a_1 = 1, g_1 = 1/v_1, c_1 = 0, e_1 = 0$, we reduce to the case where

$$H_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Next, consider the case where rank $D = 1$. Then

$$H_2 = S_1^T H_1 S_1 = \begin{bmatrix} 0 & a_1^2 & 0 & v_2 \\ a_1^2 & 0 & w_2 & 0 \\ 0 & w_2 & 0 & c_1^2 \\ v_2 & 0 & c_1^2 & 0 \end{bmatrix}$$

where

$$v_2 = c_1 a_1 + (a_1 g_1 - c_1 e_1) v_1,$$
$$w_2 = c_1 a_1 - (a_1 g_1 - c_1 e_1) v_1.$$

Taking again $a_1 = 1, c_1 = 0, e_1 = 0$ and $g_1 = 1/v_1$ we arrive at

$$H_2 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

For the case rank $D = 2$ there is no further reduction we perform. So, we arrive at the following possible simple forms: the pair (A, H) can be reduced by similarity on A and simultaneous congruence on H to a pair where $A = J_2(1) \oplus J_2(1)$, while H is of the form

$$H = \begin{cases} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} & \text{if rank } D = 0, \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} & \text{if rank } D = 1, \\ \begin{bmatrix} 0 & 1 & 0 & v_1 \\ 1 & 0 & -v_1 & 0 \\ 0 & -v_1 & 0 & 1 \\ v_1 & 0 & 1 & 0 \end{bmatrix} & \text{if rank } D = 2. \end{cases}$$

3.3.3 More Jordan blocks with equal size

In this section we discuss what happens if the matrix A has real Jordan blocks which are of equal size. First consider the case when they are of odd size. For example, let $A = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix}$ be in real Jordan form with respect to the eigenvalue $\lambda = \alpha \pm i\beta$ and where

$$J_p = \begin{bmatrix} \gamma & I & & & \\ & \gamma & I & & \\ & & \ddots & \ddots & \\ & & & \gamma & I \\ & & & & \gamma \end{bmatrix},$$

for $p = 1, 2$. Each block J_p is a $n \times n$ block matrix with $\gamma = \alpha I + \beta H_0$. Furthermore, let $H = [H_{i,j}]_{i,j=1}^2$, with each $H_{i,j}$ a $n \times n$ block matrix and $H = H^T$. If A is H-expansive we have as before that

$$\begin{aligned} 0 &\leq A^T H A - H \\ &= \begin{bmatrix} J_1^T H_{1,1} J_1 - H_{1,1} & J_1^T H_{1,2} J_2 - H_{1,2} \\ J_2^T H_{2,1} J_1 - H_{2,1} & J_2^T H_{2,2} J_2 - H_{2,2} \end{bmatrix} \end{aligned}$$

and hence for $p = 1, 2$ it must hold that $J_p^T H_{p,p} J_p - H_{p,p} \geq 0$. The form of $H_{p,p}$ is known when n is odd, see (3.18). We know that the block $J_2^T H_{2,1} J_1 - H_{2,1}$ is square and that

$$\begin{aligned} &(J_2^T H_{2,1} J_1 - H_{2,1})_{i,j} = 0 \\ \text{for} \quad &i = n+1, \ldots, 2n \quad \text{and} \quad j = 1, \ldots, [\frac{n+1}{2}] \\ \text{or} \quad &j = 1, \ldots, n \quad \text{and} \quad i = n+1, \ldots, n + [\frac{n+1}{2}]. \end{aligned}$$

As in the previous section, we let $H_{2,1} = [\tilde{H}_{i,j}]$. For convenience, recall from (3.31) that

$$(J_2^T H_{2,1} J_1 - H_{2,1})_{i,j} = \tilde{H}_{i-1,j-1} + \tilde{H}_{i-1,j}\gamma + \gamma^T \tilde{H}_{i,j-1} + \gamma^T \tilde{H}_{i,j}\gamma - \tilde{H}_{i,j}. \tag{3.33}$$

From Lemma 3.2.6 the entries on the main diagonal of $H_{2,1}$ will be of the form $aI + bH_0$ for some a and b. As before Lemma 3.2.6, Lemma 3.2.7 and Lemma 3.2.8 can be applied to the zero block entries of (3.33). If for example we let $n = 5$, then we have

$$\tilde{H}_{1,5} = \Psi^T \tilde{H}_{2,4}, \;\; \tilde{H}_{3,3} = a_1 I + b_1 H_0, \;\; \tilde{H}_{2,4} = \Psi^T \tilde{H}_{3,3}.$$

Thus, $\tilde{H}_{1,5} = (\Psi^2)^T \tilde{H}_{3,3} = (\Psi^2)^T \mathcal{H}$, where $\tilde{H}_{3,3} = \mathcal{H}$, say. (If the matrix $H_{2,1}$ would have been symmetric, then $\mathcal{H} = cI$ for some c.) So, in general we have

$$H = \left[\begin{array}{ccc|ccc} 0 & & \iddots & 0 & & \iddots \\ & c_1 I & & & \mathcal{H}^T & \\ \iddots & & * & \iddots & & * \\ \hline 0 & & \iddots & 0 & & \iddots \\ & \mathcal{H} & & & c_2 I & \\ \iddots & & * & \iddots & & * \end{array}\right].$$

It is possible to simplify H even further. By choosing the central block entries of H (the ones explicitly written out in the above form) we form a matrix B as follows:

$$B = \left[\begin{array}{cc} c_1 I & \mathcal{H}^T \\ \mathcal{H} & c_2 I \end{array}\right] = \left[\begin{array}{cc|cc} c_1 & 0 & a_1 & -b_1 \\ 0 & c_1 & b_1 & a_1 \\ \hline a_1 & b_1 & c_2 & 0 \\ -b_1 & a_1 & 0 & c_2 \end{array}\right].$$

(If $H_{2,1}$ is symmetric then we can omit all zero entries and consider $B = \left[\begin{array}{cc} c_1 & a_1 \\ a_1 & c_2 \end{array}\right]$.)

Since $B = B^T$, there exists an orthogonal matrix P and a diagonal matrix D with the eigenvalues of B on its main diagonal such that

$$B = P^{-1} D P.$$

Because of the special structure of B, we actually obtain something stronger:

$$\left[\begin{array}{c} u \\ v \end{array}\right] \in \mathrm{Ker}(B - \lambda I_4) \Leftrightarrow \left[\begin{array}{c} H_0 u \\ H_0 v \end{array}\right] \in \mathrm{Ker}(B - \lambda I_4).$$

Furthermore, the set $\{\begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} H_0u \\ H_0v \end{bmatrix}\}$ with either u or v nonzero, is linearly independent. Thus B has at most two different eigenvalues which we denote by $\tilde{h}_1$ and $\tilde{h}_2$, respectively. Therefore, if $\begin{bmatrix} u \\ v \end{bmatrix}$ is an eigenvector of B corresponding to $\tilde{h}_1$ and $\begin{bmatrix} u' \\ v' \end{bmatrix}$ is an eigenvector of B corresponding to $\tilde{h}_2$, then the set of vectors $\{\begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} H_0u \\ H_0v \end{bmatrix}, \begin{bmatrix} u' \\ v' \end{bmatrix} \begin{bmatrix} H_0u' \\ H_0v' \end{bmatrix}\}$ is linearly independent. Thus, let

$$P = \left[\begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} H_0u \\ H_0v \end{bmatrix}, \begin{bmatrix} u' \\ v' \end{bmatrix}, \begin{bmatrix} H_0u' \\ H_0v' \end{bmatrix} \right]$$

$$D = \text{diag}\{\tilde{h}_1, \tilde{h}_1, \tilde{h}_2, \tilde{h}_2\},$$

with $\tilde{h}_1, \tilde{h}_2$ the eigenvalues of B.

Alternatively, we can describe the diagonalization of the matrix B directly in terms of the Jordan basis. Let $x_1, \ldots, x_{2n}$ and $y_1, \ldots, y_{2n}$ be two Jordan chains corresponding to the two Jordan blocks J_1 and J_2 respectively. Then we have for the matrix B that:

$$\begin{aligned}
c_1 &= \langle Hx_n, x_n\rangle = \langle Hx_{n+1}, x_{n+1}\rangle, \\
c_2 &= \langle Hy_n, y_n\rangle = \langle Hy_{n+1}, y_{n+1}\rangle, \\
a_1 &= \langle Hx_n, y_n\rangle = \langle Hy_n, x_n\rangle = \langle Hx_{n+1}, y_{n+1}\rangle = \langle Hy_{n+1}, x_{n+1}\rangle, \\
b_1 &= \langle Hx_{n+1}, y_n\rangle = \langle Hy_n, x_{n+1}\rangle, \\
-b_1 &= \langle Hx_n, y_{n+1}\rangle = \langle Hy_{n+1}, x_n\rangle.
\end{aligned}$$

It is easy to verify that any linear combination of $x_1, \ldots, x_{2n}$ and $y_1, \ldots, y_{2n}$ is again a Jordan chain. Therefore, consider the new Jordan chains given by

$$\begin{aligned}
\tilde{x}_i &= f_1x_i + d_1y_i \\
\tilde{y}_i &= f_2x_i + d_2y_i \quad \text{for} \quad i = 1, \ldots, 2n,
\end{aligned}$$

where $f_1, f_2, d_1, d_2 \in \mathbb{R}$.

In terms of the new Jordan basis, the diagonalization of matrix B is

$$\widetilde{B} = \begin{bmatrix}
\langle H\tilde{x}_n, \tilde{x}_n\rangle & \langle H\tilde{x}_n, \tilde{x}_{n+1}\rangle & \langle H\tilde{x}_n, \tilde{y}_n\rangle & \langle H\tilde{x}_n, \tilde{y}_{n+1}\rangle \\
\langle H\tilde{x}_{n+1}, \tilde{x}_n\rangle & \langle H\tilde{x}_{n+1}, \tilde{x}_{n+1}\rangle & \langle H\tilde{x}_{n+1}, \tilde{y}_n\rangle & \langle H\tilde{x}_{n+1}, \tilde{y}_{n+1}\rangle \\
\langle H\tilde{x}_n, \tilde{y}_n\rangle & \langle H\tilde{x}_{n+1}, \tilde{y}_n\rangle & \langle H\tilde{y}_n, \tilde{y}_n\rangle & \langle H\tilde{y}_n, \tilde{y}_{n+1}\rangle \\
\langle H\tilde{x}_n, \tilde{y}_{n+1}\rangle & \langle H\tilde{x}_{n+1}, \tilde{y}_{n+1}\rangle & \langle H\tilde{y}_{n+1}, \tilde{y}_n\rangle & \langle H\tilde{y}_{n+1}, \tilde{y}_{n+1}\rangle
\end{bmatrix},$$

with

$$\langle H\tilde{x}_i, \tilde{x}_i\rangle \neq 0, \qquad \langle H\tilde{y}_i, \tilde{y}_i\rangle \neq 0 \quad \text{for} \quad i = n, n+1$$

and

$$\langle H\tilde{x}_i, \tilde{x}_j\rangle = 0, \qquad \langle H\tilde{y}_i, \tilde{y}_j\rangle = 0 \quad \text{for} \quad i \neq j,$$

$$\langle H\tilde{x}_i, \tilde{y}_i\rangle = 0, \qquad \langle H\tilde{x}_i, \tilde{y}_j\rangle = 0 \quad \text{for all} \quad i,j = n, n+1.$$

Thus, under these conditions $\tilde{B}$ is diagonal and one can solve for $f_1, f_2, d_1, d_2 \in \mathbb{R}$.

The above diagonalizing procedure (for the matrix B) produces two Jordan chains $\{\tilde{x}_1, \ldots, \tilde{x}_{2n}\}$ and $\{\tilde{y}_1, \ldots, \tilde{y}_{2n}\}$ with respect to which H reduces to the following form:

$$\tilde{H} = \left[\begin{array}{cccc|cccc} 0 & & & \iddots & 0 & & & 0 \\ & & \tilde{h}_1 I & & & & \iddots & \iddots \\ & \iddots & & & & \iddots & \mathcal{H}^T & \\ \iddots & & & * & 0 & \iddots & & * \\ \hline 0 & & & 0 & 0 & & & \iddots \\ & & \iddots & \iddots & & & \tilde{h}_2 I & \\ & \iddots & \mathcal{H} & & & \iddots & & \\ 0 & \iddots & & * & \iddots & & & * \end{array}\right].$$

Suppose that A consists of k Jordan blocks on the diagonal, say each of the same odd size n, corresponding to only a single pair of eigenvalues $\lambda = \alpha \pm i\beta$. Let each J_p for $p = 1, \ldots, k$ be defined as before. Now, let $H = [H_{i,j}]_{i,j=1}^{k}$ and $H = H^T$. If A is H-expansive we can again apply Lemma 3.2.6, Lemma 3.2.7 and Lemma 3.2.8 as before to find the general form for H, say

$$H = \left[\begin{array}{ccc|ccc|c|cc} 0 & & \iddots & 0 & & \iddots & & & \\ & C_{11} & & & C_{12} & & & & \\ \iddots & & * & \iddots & & * & & & \\ \hline 0 & & \iddots & 0 & & \iddots & & & \\ & C_{12} & & & C_{22} & & & & \\ \iddots & & * & \iddots & & * & & & \\ \hline \vdots & & & & & & \ddots & & \\ \hline & & & & & & & 0 & \iddots \\ & & & & & & & & C_{kk} \\ & & & & & & & \iddots & * \end{array}\right]. \tag{3.34}$$

Now we pick out all the terms in the middle of each small block to construct B, that is

$$B = \begin{bmatrix} C_{11} & C_{12} & \ldots & C_{1k} \\ C_{12} & C_{22} & \ldots & C_{2k} \\ \vdots & & \ddots & \\ C_{1k} & \ldots & \ldots & C_{kk} \end{bmatrix},$$

with $C_{kk} = c_k I$ and $C_{kl} = a_{k,l} I + b_{k,l} H_0$ for some $a_{k,l}$ and $b_{k,l}$, when $k \neq l$. Note that B is symmetric, hence $C_{k,l} = C_{l,k}$. After diagonalization, B will be similar to D

$$D = \operatorname{diag}\{\tilde{h}_1, \tilde{h}_1, \ldots, \tilde{h}_k, \tilde{h}_k\},$$

with $\tilde{h}_1, \ldots, \tilde{h}_k$ the eigenvalues of B and

$$P = \left[\begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \begin{bmatrix} H_0 u_1 \\ H_0 v_1 \end{bmatrix}, \ \ldots, \ \begin{bmatrix} u_k \\ v_k \end{bmatrix}, \begin{bmatrix} H_0 u_k \\ H_0 v_k \end{bmatrix} \right],$$

where $B = P^{-1} D P$. For $i = 1, \ldots, k$ we have that

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} \in \mathrm{Ker}\{\begin{bmatrix} C_{i,i} & C_{i,i+1} \\ C_{i+1,i} & C_{i+1,i+1} \end{bmatrix} - \lambda_i I_4\}.$$

With respect to the new Jordan basis as explained above, H will be simplified to

$$\widetilde{H} = \left[\begin{array}{ccc|ccc|c|c} 0 & & \iddots & 0 & & 0 & & \\ & \tilde{C}_{11} & & & \iddots & & & \\ \iddots & & * & 0 & & * & & \\ \hline 0 & & 0 & 0 & & \iddots & & \\ & \iddots & & & \tilde{C}_{22} & & & \\ 0 & & * & \iddots & & * & & \\ \hline \vdots & & & & & & \ddots & \\ \hline & & & & & & 0 & \iddots \\ & & & & & & & \tilde{C}_{kk} \\ & & & & & & \iddots & * \end{array}\right]. \tag{3.35}$$

This form of $\tilde{H}$ will give us better information on the vectors to be selected from the newly generated Jordan basis for the construction of the invariant nonnegative and nonpositive subspaces.

A similar situation occurs in the special case when $\lambda = 1$ or $\lambda = -1$. Because we are interested in constructing a simple form from which one can easily read off invariant maximal semidefinite subspaces, we only need to look into the case where the sizes of the Jordan blocks are odd. Note that this avoids the difficulties with blocks of even size which were illustrated at the end of the previous subsection. Nothing changes in the general form of the matrix H, see (3.34). The only difference again is that the entries are scalar entries and not 2×2 block entries. The diagonalization process as described above can be performed when we let B consist of all the central entries of H. The newly formed matrix $\tilde{H}$ (similar to (3.35)) has the correct signs for the central entries on the main diagonal.

We illustrate the procedure by the following example: Suppose we have obtained the following 6×6 matrix H:

$$H = \left[\begin{array}{ccc|ccc} 0 & 0 & * & 0 & 0 & * \\ 0 & 1 & * & 0 & 2 & * \\ * & * & * & * & * & * \\ \hline 0 & 0 & * & 0 & 0 & * \\ 0 & 2 & * & 0 & 1 & * \\ * & * & * & * & * & * \end{array}\right]. \tag{3.36}$$

Recall that:

$$H = \left[\begin{array}{ccc|ccc} [x_1,x_1] & [x_2,x_1] & [x_3,x_1] & [y_1,x_1] & [y_2,x_1] & [y_3,x_1] \\ [x_1,x_2] & [x_2,x_2] & [x_3,x_2] & [y_1,x_2] & [y_2,x_2] & [y_3,x_2] \\ [x_1,x_3] & [x_2,x_3] & [x_3,x_3] & [y_1,x_3] & [y_2,x_3] & [y_3,x_3] \\ \hline [x_1,y_1] & [x_2,y_1] & [x_3,y_1] & [y_1,y_1] & [y_2,y_1] & [y_3,y_1] \\ [x_1,y_2] & [x_2,y_2] & [x_3,y_2] & [y_1,y_2] & [y_2,y_2] & [y_3,y_2] \\ [x_1,y_3] & [x_2,y_3] & [x_3,y_3] & [y_1,y_3] & [y_2,y_3] & [y_3,y_3] \end{array}\right],$$

where x_1, x_2, x_3 and y_1, y_2, y_3 are Jordan chains associated with the blocks J_1, J_2 of the eigenvalue $\lambda = \alpha \pm i\beta$.

So, we have $1 = \langle Hx_2, x_2\rangle = \langle Hy_2, y_2\rangle$ and $2 = \langle Hx_2, y_2\rangle = \langle Hy_2, x_2\rangle$. According to our discussion above, we take $B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Using the standard diagonalization procedure, there exist a diagonal matrix D and an orthogonal matrix P such that $B = P^{-1}DP$ and where

$$D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}; P = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

We take the columns of P as coefficients in the following linear combinations,

$$\begin{aligned} \tilde{x}_i &= x_i + y_i \\ \tilde{y}_i &= x_i - y_i \quad \text{for} \quad i = 1,2,3 \end{aligned}$$

of the two Jordan chains. So, we obtain two new Jordan chains $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ and $\tilde{y}_1, \tilde{y}_2, \tilde{y}_3$ corresponding to $\lambda = \alpha \pm i\beta$ with respect to the Jordan blocks J_1, J_2. In terms of the two new Jordan chains we obtain a simplified form of H, given by

$$\tilde{H} = \left[\begin{array}{ccc|ccc} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 6 & * & 0 & 0 & * \\ * & * & * & 0 & * & * \\ \hline 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & -2 & * \\ 0 & * & * & * & * & * \end{array}\right], \tag{3.37}$$

where we see that

$$\langle H\tilde{x}_2, \tilde{x}_2\rangle = 6, \quad \langle H\tilde{y}_2, \tilde{x}_2\rangle = 0, \quad \langle H\tilde{x}_2, \tilde{y}_2\rangle = 0, \quad \langle H\tilde{y}_2, \tilde{y}_2\rangle = -2.$$

3.3.4 More Jordan blocks: different eigenvalues on the unit circle

In this section we consider several Jordan blocks with different eigenvalues. We once again illustrate this case by an example, after which we will give the simple form for the matrix H.

Let $A = J_1 \oplus J_2$ with J_i a real Jordan block of size $n_i \times n_i$ corresponding to the different pairs of eigenvalues $\lambda_1, \bar{\lambda}_1$ and $\lambda_2, \bar{\lambda}_2$, respectively. In other words we have $\sigma(J_1) \cap \sigma(J_2) = \emptyset$ and $\sigma(J_i) \subseteq \mathbb{T}$. Let $H = [H_{i,j}]_{i,j=1}^2$ and let A be H-expansive, then as before

$$0 \leq A^T H A - A = \begin{bmatrix} J_1^T H_{1,1} J_1 - H_{1,1} & J_1^T H_{1,2} J_2 - H_{1,2} \\ J_2^T H_{2,1} J_1 - H_{2,1} & J_2^T H_{2,2} J_2 - H_{2,2} \end{bmatrix}.$$

The two blocks on the main diagonal can be treated as before. The difference in this case lies in the other two blocks, i.e., $J_1^T H_{1,2} J_2 - H_{1,2}$ and $J_2^T H_{2,1} J_1 - H_{2,1}$. We know that

$$\begin{aligned} (J_1^T H_{1,1} J_1 - H_{1,1})_{i,j} = 0 \quad \text{for} \quad 1 \leq i \leq [\frac{n_1+1}{2}] \quad \text{or} \quad 1 \leq j \leq [\frac{n_1+1}{2}], \\ (J_2^T H_{2,2} J_2 - H_{2,2})_{i,j} = 0 \quad \text{for} \quad 1 \leq i \leq [\frac{n_2+1}{2}] \quad \text{or} \quad 1 \leq j \leq [\frac{n_2+1}{2}]. \end{aligned}$$

This means that in addition

$$\begin{aligned} & (J_2^T H_{2,1} J_1 - H_{2,1})_{i,j} = 0 \\ \text{for} \quad & i = n_1+1, \ldots, n_1 + [\frac{n_2+1}{2}] \quad \text{and} \quad j = 1, \ldots, n_1, \\ \text{or} \quad & j = 1, \ldots, [\frac{n_1+1}{2}] \quad \text{and} \quad i = n_1+1, \ldots, n_1+n_2. \end{aligned}$$

Let

$$B = J_2^T H_{2,1} J_1 - H_{2,1}. \tag{3.38}$$

We have that J_2^T is of size $n_2 \times n_2$, J_1 is of size $n_1 \times n_1$ and B (in equation 3.38) is a $n_2 \times n_1$ matrix with 2×2 blocks as entries. It is also known that $B_{i,j} = 0$ for

$$\begin{aligned} & i = n_1+1, \ldots, n_1 + [\frac{n_2+1}{2}] \quad \text{and} \quad j = 1, \ldots, n_1 \\ \text{or} \quad & j = 1, \ldots, [\frac{n_1+1}{2}] \quad \text{and} \quad i = n_1+1, \ldots, n_1+n_2. \end{aligned}$$

The remaining part of our discussion will make use of the following facts:

Fact 1: Let $C = (c_{ij})$ be a $r \times s$ (scalar) matrix. Then $c_{ij} = 0$ for $j = 1, \ldots, [\frac{s+1}{2}]$ and $i = 1, \ldots, r$ if and only if $CP = 0$, where P is the $s \times [\frac{s+1}{2}]$ (scalar) matrix

$$P = \begin{bmatrix} I_{[\frac{s+1}{2}]} \\ 0 \end{bmatrix}.$$

Fact 2: Let $C = (c_{ij})$ be a $r \times s$ (scalar) matrix. Then $c_{ij} = 0$ for $i = 1, \ldots, [\frac{r+1}{2}]$ and $j = 1, \ldots, s$ if and only if $QC = 0$, where Q is the $[\frac{r+1}{2}] \times r$ (scalar) matrix

$$Q = \begin{bmatrix} I_{2[\frac{r+1}{2}]} & 0 \end{bmatrix}.$$

We state here Theorem A.4.1 in [24].

Theorem 3.3.2 *Let A and B be (real or complex) matrices of sizes $m \times m$ and $n \times n$, respectively. Then the matrix equation*

$$AX - XB = 0$$

with the unknown (real or complex) $m \times n$ matrix X has only the trivial solution $X = 0$ if and only if $\sigma(A) \cap \sigma(B) = \emptyset$.

From the discussion above it is therefore clear that both $BP = 0$ and $QB = 0$ hold, where in this case we let P be the $n_1 \times [\frac{n_1+1}{2}]$ block matrix

$$P = \begin{bmatrix} I_{[\frac{n_1+1}{2}]} \\ 0 \end{bmatrix}.$$

with 2×2 blocks as entries (i.e. P is a $2n_1 \times 2[\frac{n_1+1}{2}]$ scalar matrix) and Q the $[\frac{n_2+1}{2}] \times n_2$ block matrix

$$Q = \begin{bmatrix} I_{2[\frac{n_2+1}{2}]} & 0 \end{bmatrix}.$$

with 2×2 blocks as entries (i.e. Q is a $2[\frac{n_2+1}{2}] \times 2n_2$ scalar matrix).

Multiplying (3.38) from the right by P and using Fact 1, we therefore have

$$J_2^T H_{2,1} J_1 P - H_{2,1} P = BP = 0. \tag{3.39}$$

Consider

$$J_1 P = \begin{bmatrix} \gamma & I & & \\ & \gamma & I & \\ & & \ddots & I \\ & & & \gamma \end{bmatrix} \begin{bmatrix} I_{[\frac{n_1+1}{2}]} \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma & I & \\ & \ddots & I \\ & & \gamma \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \tilde{J}_1 \\ 0 \end{bmatrix}$$

where γ is the 2×2 block associated with the eigenvalue λ_1 and where $\tilde{J}_1$ is a $[\frac{n_1+1}{2}]\times[\frac{n_1+1}{2}]$ block.

We write $H_{2,1}$ in the following block form:

$$H_{2,1} = \begin{bmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{bmatrix},$$

where each $X_{i,j}$ has the suitable dimension, such that the following computation can be done:

$$H_{2,1} J_1 P = \begin{bmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{bmatrix} \begin{bmatrix} \tilde{J}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} X_{1,1} \\ X_{2,1} \end{bmatrix} \tilde{J}_1,$$

which is a $n_2 \times [\frac{n_1+1}{2}]$ block matrix. Also,

$$H_{2,1}P = \begin{bmatrix} X_{1,1} \\ X_{2,1} \end{bmatrix}$$

which is a $n_2 \times [\frac{n_1+1}{2}]$ block matrix. Hence from (3.39) we have

$$J_2^T(H_{2,1}P)\tilde{J}_1 - H_{2,1}P = 0.$$

By our assumption that $\sigma(J_1) \cap \sigma(J_2) = \emptyset$ and $\sigma(J_i) \subseteq \mathbb{T}$ we find that $\sigma(J_2^T) \cap \sigma(\tilde{J}_1^{-1}) = \emptyset$, where $\sigma(\tilde{J}_1) = \sigma(J_1)$. Then by Theorem 3.3.2 it follows that $H_{2,1}P = 0$.

Since $QB = 0$ from Fact 2, if we multiply (3.38) from the left by Q and then take the transpose, we get

$$J_1^T H_{2,1}^T J_2 Q^T - H_{2,1}^T Q^T = 0.$$

Following the above arguments, we may therefore also show that

$$J_1(H_{2,1}^T Q^T)\tilde{J}_2 - (H_{2,1}^T Q^T) = 0.$$

for the $[\frac{n_2+1}{2}] \times [\frac{n_2+1}{2}]$ block matrix

$$\tilde{J}_2 = \begin{bmatrix} \gamma & I & & \\ & \gamma & I & \\ & & \ddots & I \\ & & & \gamma \end{bmatrix}$$

such that

$$J_2 Q^T = \begin{bmatrix} \tilde{J}_2 \\ 0 \end{bmatrix}.$$

Again by Theorem 3.3.2 it then follows that $H_{2,1}^T Q^T = 0$ and hence that $QH_{2,1} = 0$.

Thus, in our case with a more general J_1 and J_2 we have that $H_{i,j} = 0$ for the same entries for which $(J_2^T H_{2,1} J_1 - H_{2,1})_{i,j} = 0$. For example, if $n_1 = 6$ and $n_2 = 4$, we have that

$$H = \left[\begin{array}{cccccc|cccc} 0 & 0 & 0 & 0 & 0 & \iddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \iddots & & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \iddots & & & 0 & 0 & 0 & 0 \\ 0 & 0 & \iddots & & * & & 0 & 0 & & \\ 0 & \iddots & & & & & 0 & 0 & * & \\ \iddots & & & & & & 0 & 0 & & \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \iddots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \iddots & * \\ 0 & 0 & 0 & * & * & * & 0 & \iddots & * & * \\ 0 & 0 & 0 & * & * & * & \iddots & * & * & * \end{array}\right]. \tag{3.40}$$

Also in this section we briefly want to consider the special case when $\lambda = \pm 1$. So, if we let $A = J_1 \oplus J_2$, with J_i a Jordan block of size $n_i \times n_i$ for $i = 1, 2$ corresponding to the

eigenvalues $\lambda = 1$ and $\lambda = -1$, respectively or vice versa. It may also happen that one eigenvalue is $\lambda = \pm 1$ and the other a complex eigenvalue $\lambda = \alpha \pm i\beta$ on the unit circle. In this case the one Jordan block consists of scalar entries and the other Jordan block of 2×2 entries. As a convention, we will always put the scalar case in the (1,1) position.

By choosing the dimensions of P and Q wisely, we obtain similarly that $H_{ij} = 0$ for the same entries for which $(J_2^T H_{21} J_1 - H_{21})_{ij} = 0$. So, if we have any of the above mentioned scenarios, we still obtain a matrix H similar to what we have in (3.40). If we have eigenvalue $\lambda = \pm 1$ and eigenvalue $\lambda = \alpha \pm i\beta$, then the entries in the (1,1)-block of H are scalar entries, the (2,1)-block has 2×1 entries, the (1,2)-block has 1×2 entries and the (2,2)-block has 2×2 entries.

3.4 The general case

In this section we will consider the most general form of the matrix A. This is where A can consist of blocks which fall into the different cases as discussed before. Furthermore we will also include the eigenvalues which do not belong to the unit circle. Therefore, let

$$A = \begin{bmatrix} J_1 & & & & & & \\ & J_2 & & & & & \\ & & J_3 & & & & \\ & & & \ddots & & & \\ & & & & J_p & & \\ & & & & & J_{p+1} & \\ & & & & & & J_{p+2} \end{bmatrix} \quad \text{and} \quad H = \left[\; H_{i,j} \; \right]_{i,j=1}^{p+2},$$

where in A we have

$$\begin{aligned} &\sigma(J_1) = \{1\}, \\ &\sigma(J_2) = \{-1\}, \\ &\sigma(J_j) = \{\alpha_j \pm i\beta_j\}, \alpha_j^2 + \beta_j^2 = 1, \beta_j \neq 0 \quad \text{for} \quad j = 3, \ldots, p, \\ &\sigma(J_{p+1}) \subset \mathbb{D} = \{\lambda \mid |\lambda| < 1\}, \\ &\sigma(J_{p+2}) \subset \mathbb{D}^e = \{\lambda \mid |\lambda| > 1\}, \end{aligned}$$

and $\sigma(J_i) \cap \sigma(J_j) = \emptyset$ for $i \neq j$, $i, j = 3, \ldots, p$. We assume that $H_{i,j}$ is as before.

In this case, J_1 and J_2 are Jordan matrices with standard Jordan blocks with eigenvalues 1 and -1, respectively, J_i for $i = 3, \ldots, p$ are in real canonical form and J_{p+1}, J_{p+2} are Jordan matrices with blocks that are either in real Jordan canonical form (in case they have non-real eigenvalues) or in standard Jordan canonical form (in case they have real eigenvalues). The form of the corresponding matrix H in the odd case is given by (3.29) and in the even case by (3.30).

Let

$$J_j = \begin{bmatrix} J_{j_1} & & \\ & \ddots & \\ & & J_{j_{n_j}} \end{bmatrix},$$

for $j = 3, \ldots p$. Here, each J_{j_i} is an $n_i \times n_i$ real Jordan block belonging to an eigenvalue λ_j with its corresponding $\gamma = \alpha I + \beta H_0$ and $\beta \neq 0$. Here n_i (the number of 2×2-blocks) can be odd or even. In the odd case the corresponding matrix H is of the form given by (3.18) and in the even case H is of the form (3.19). For $j = p+1, p+2$ no structure of H is required.

Theorem 3.4.1 *Let $H = H^T$ be an $n \times n$ real symmetric invertible matrix, and let A be H-expansive, i.e. $A^T H A - H \geq 0$. Then there is a real matrix S such that $S^{-1}AS$ is in real Jordan form $S^{-1}AS = J_1 \oplus J_2 \oplus J_3 \oplus \ldots \oplus J_p \oplus J_{p+1} \oplus J_{p+2}$, and simultaneously $S^T H S = (H_{ij})_{i,j=1}^{p+2}$ where*

$$H_{ij} = (H_{k,l}^{(i,j)})_{k,l=1}^{n_j}.$$

Here the blocks J_j and $H_{k,l}^{(i,j)}$ satisfy the following:

1. $$\begin{aligned} J_1 &= J_{1,1}(1) \oplus \ldots \oplus J_{1,n_1}(1) \\ J_2 &= J_{2,1}(-1) \oplus \ldots \oplus J_{2,n_2}(-1), \end{aligned}$$

   where each $J_{1,j}$, respectively, $J_{2,j}$ is a Jordan block. The block $H_{k,k}^{(1,1)}$ has the same size as $J_{j,k}$ for $k = 1, \ldots, n_j$, then for $i = 1, 2$

   $$\begin{aligned} &H_{k,k}^{(i,i)} \quad \textit{is as in } (3.29) \quad \textit{if size of } J_{j,k} \textit{ is odd} \\ &H_{k,k}^{(i,i)} \quad \textit{is as in } (3.30) \quad \textit{if size of } J_{j,k} \textit{ is even, possibly with } c = 0 \\ &H_{k,l}^{(i,i)} \quad \textit{for } k \neq l \textit{ some lower anti-diagonal matrix.} \end{aligned}$$

2. *For $k = 3, \ldots, p$, J_{k,n_k} is a real Jordan matrix with eigenvalues $\alpha_k \pm i\beta_k$, $\beta_k \neq 0$ and $\alpha_k^2 + \beta_k^2 = 1$. The corresponding simplified form for the matrix $H^{(j,j)}$ is:*

   as in (3.18) if size J_k is odd and $n_j = 1$
   as in (3.19) if size J_k is even and $n_j = 1$
   as in (3.32) if $n_j > 1$
   as in (3.35) if size J_k is odd and contains equal size blocks

   (See the end of Sections 3.3.2, 3.3.3 and 3.3.4 if $\beta = 0$.)

3. *Blocks $H_{k,l}^{(i,j)}$ with $i \neq j$ are as in* (3.40).

4. *For $k = p+1, p+2$, J_k is a real matrix in real Jordan canonical form with $\sigma(J_{p+1})$ in the open unit disk and $\sigma(J_{p+2})$ in the exterior of the closed unit disk.*

This simplification of H is required for the construction of maximal invariant subspaces of a H-expansive matrix A in Section 3.5.

3.5 Construction of invariant subspaces

In this section we will describe how the invariant subspaces are constructed using the simple form of H. At the end we will also show that these subspaces are indeed maximal.

Let $\sigma(A)\cap\mathbb{T} = \{\lambda_1, \ldots, \lambda_p\}$, where we order the eigenvalues as follows: $\lambda_1 = 1, \lambda_2 = -1$ (if either or both are in the spectrum of A), $\lambda_3, \ldots, \lambda_p$ are the remaining eigenvalues of A on the unit circle, grouped so that $\lambda_{2j} = \overline{\lambda}_{2j-1}$ and $\mathrm{Im}\lambda_{2j-1} > 0$. See also Section 3.4.

Let $\mathcal{N}_j$ be the space spanned by all the eigenvectors and generalized eigenvectors corresponding to

a) the eigenvalue $\lambda_j = 1$ (in case $1 \in \sigma(A)$);

b) the eigenvalue $\lambda_j = -1$ (in case $-1 \in \sigma(A)$);

c) pairs of eigenvalues $\lambda_j = \alpha_j \pm i\beta_j$ with $\alpha_j^2 + \beta_j^2 = 1$, $\beta_j \neq 0$.

Further, let $\mathcal{N}_{p+1}$ be the subspace spanned by all eigenvectors and generalized eigenvectors corresponding to eigenvalues in $\mathbb{D}$, and likewise $\mathcal{N}_{p+2}$ the spectral subspace of A corresponding to eigenvalues in $\mathbb{D}^e$, where $\mathbb{D}^e = \{\lambda \mid |\lambda| > 1\}$.

We have to distinguish between the nonpositive and nonnegative invariant subspaces, say $\mathcal{M}_+$ and $\mathcal{M}_-$. So, if A is H-expansive then

$$0 \leq \langle (A^T H A - H)x, x\rangle = (|\lambda|^2 - 1)\langle Hx, x\rangle,$$

for every eigenvalue λ of A and corresponding eigenvector x. Thus, if $\lambda \in \mathbb{D}$, then $\langle Hx, x\rangle \leq 0$ for all $x \in \mathcal{N}_{p+1}$. Therefore $\mathcal{N}_{p+1}$ is an H-nonpositive subspace. If $\lambda \in \mathbb{D}^e$ then $\langle Hx, x\rangle \geq 0$ for all $x \in \mathcal{N}_{p+2}$. Therefore $\mathcal{N}_{p+2}$ is an H-nonnegative subspace.

For $|\lambda_j| = 1$ we define $\mathcal{N}_j^+$ and $\mathcal{N}_j^-$ as follows. We assume, as usual, that A is in real Jordan canonical form, and that H is in the simple form as constructed in the previous sections. In particular, with respect to odd size blocks we assume the construction at the end of Section 3.3 to be carried out. Let

$$J_j = \begin{bmatrix} J_{j1} & & \\ & \ddots & \\ & & J_{jnj} \end{bmatrix}.$$

as in Section 3.4. Denote by k_{ji} the size of the block J_{ji} if $j = 1, 2$. Corresponding to each block construct $\mathcal{N}_{ji}$ as follows. Suppose $x_{ji1}, \ldots, x_{jik_{ji}}$ for $j = 1, 2$, are the corresponding basis vectors from a real Jordan basis for A. Then set

$$\mathcal{N}_{ji}^{\pm} = \mathrm{span}\{x_{ji1}, \ldots, x_{jik_{ji}/2}\}$$

when k_{ji} is even, and set

$$\mathcal{N}_{ji}^{\pm} = \mathrm{span}\{x_{ji1}, \ldots, x_{ji\frac{k_{ji}+1}{2}}\}$$

when k_{ji} is odd and the corresponding sign is $+1$, (and set this to be $\mathcal{N}_{ji}^-$ when the sign is -1). Put

$$\mathcal{N}_{ji}^{+} = \mathrm{span}\{x_{ji1}, \ldots, x_{ji\frac{k_{ji}-1}{2}}\}$$

when k_{ji} is odd and the corresponding sign is -1 (and set this to be $\mathcal{N}_{ji}^-$ when the sign is $+1$). Then, for $j = 1, 2$ set

$$\mathcal{N}_j^{\pm} = \bigoplus_{i=1}^{n_j} \mathcal{N}_{ji}^{\pm}.$$

For $j = 3, \ldots, p$, denote by $2k_{ji}$ the size of the block J_{ji}. Suppose $x_{ji1}, \ldots, x_{ji2k_{ji}}$ are the corresponding basis vectors from a real Jordan basis for A, then set

$$\mathcal{N}_j^{\pm} = \operatorname{span}\{x_{ji1}, \ldots, x_{jik_{ji}}\}$$

when k_{ji} is even, and set

$$\mathcal{N}_{ji}^{+} = \operatorname{span}\{x_{ji1}, \ldots, x_{jik_{ji}+1}\}$$

when k_{ji} is odd and the corresponding sign $\tilde{c}$ on the center position of the corresponding block in the form $\tilde{H}$ (see the end of Section 3.3.3) is positive, (and set this to be $\mathcal{N}_{ji}^-$ if the sign is negative), set

$$\mathcal{N}_{ji}^{+} = \operatorname{span}\{x_{ji1}, \ldots, x_{jik_{ji}-1}\}$$

when k_{ji} is odd and the corresponding sign is negative (and set this to $\mathcal{N}_{ji}^-$ if the sign is positive). Then, for $j = 3, \ldots, p$, set

$$\mathcal{N}_j^{\pm} = \bigoplus_{i=1}^{n_j} \mathcal{N}_{ji}^{\pm}.$$

Observe that by the construction of the simple form the spaces $\mathcal{N}_{ji}^+$ are mutually H-*orthogonal*. It is here that we use in particular the final part of Section 3.3.3. Since the direct sum of H-nonnegative subspaces which are H-orthogonal is again H-nonnegative (and likewise for H-nonpositive subspaces), we see that the $\mathcal{N}_j^+$ for $j = 1, \ldots, p$ are H-nonnegative subspaces. Now, let

$$\mathcal{M}_+ = \mathcal{N}_1^+ \oplus \ldots \oplus \mathcal{N}_p^+ \oplus \mathcal{N}_{p+2}$$

and

$$\mathcal{M}_- = \mathcal{N}_1^- \oplus \ldots \oplus \mathcal{N}_p^- \oplus \mathcal{N}_{p+1}.$$

Then $\mathcal{M}_+$ is A-invariant and H-nonnegative, while $\mathcal{M}_-$ is A-invariant and H-nonpositive. Indeed, by construction of the simple form, the spaces $\mathcal{N}_j^+$ are mutually H-orthogonal and H-orthogonal to $\mathcal{N}_{p+2}$. Hence $\mathcal{M}_+$ is H-nonnegative. Similarly, $\mathcal{M}_-$ is H-nonnegative.

The maximality of the subspaces $\mathcal{M}_+$ and $\mathcal{M}_-$ follows from a dimension argument. Given that A and H are matrices of size $n \times n$ with $\mathcal{M}_+$ a H-nonnegative subspace with dimension n_+ and $\mathcal{M}_-$ a H-nonpositive subspace with dimension n_-, we have $n = n_+ + n_-$. Then $\mathcal{M}_+$ is maximal H-nonnegative and $\mathcal{M}_-$ is maximal H-nonpositive: Indeed, from Theorem 2.3.2 in [24] we know that the dimension of a maximal H-nonnegative subspace equals $\kappa_+(H)$, the number of positive eigenvalues of H and the dimension of a H-nonpositive subspace equals $\kappa_-(H)$, the number of negative eigenvalues of H. Let $\mathcal{U}_+$ be this maximal H-nonnegative subspace with $\mathcal{M}_+ \subset \mathcal{U}_+$ ($n_+ \leq \dim \mathcal{U}_+$) and $\mathcal{U}_-$ be the maximal H-nonpositive subspace with $\mathcal{M}_- \subset \mathcal{U}_-$ ($n_- \leq \dim \mathcal{U}_-$). Now

$$\begin{aligned} n = \dim \mathcal{M}_+ + \dim \mathcal{M}_- &\leq \dim \mathcal{U}_+ + \dim \mathcal{U}_- \\ &= \kappa_+(H) + \kappa_-(H) \\ &= n \end{aligned}$$

But then $\dim \mathcal{M}_+ = \dim \mathcal{U}_+$, hence $\mathcal{M}_+ = \mathcal{U}_+$ and also $\dim \mathcal{M}_- = \dim \mathcal{U}_-$, hence $\mathcal{M}_- = \mathcal{U}_-$.

Returning to the example with H as in (3.36) we now use the form as in (3.37). We may proceed as follows to find invariant maximal nonnegative and maximal nonpositive subspaces, respectively. Since $[\tilde{x}_1, \tilde{x}_1] = [\tilde{y}_1, \tilde{y}_1] = 0$, it is clear that the vectors $\tilde{x}_1$ and $\tilde{y}_1$ belong to neutral subspaces and in particular to the neutral subspace $\mathrm{Sp}\{\tilde{x}_1, \tilde{y}_1\}$. Since $[\tilde{x}_2, \tilde{x}_2] = 6 > 0$, we may therefore conclude that

$$\mathcal{M}_+ = \mathrm{Sp}\{\tilde{x}_1, \tilde{x}_2, \tilde{y}_1\}$$

is a nonnegative subspace. Similarly, since $[\tilde{y}_2, \tilde{y}_2] = -2 < 0$, we may therefore conclude that

$$\mathcal{M}_- = \mathrm{Sp}\{\tilde{x}_1, \tilde{y}_1, \tilde{y}_2\}$$

is a nonpositive subspace.

Both $\mathcal{M}_+, \mathcal{M}_-$ being 3-dimensional mutually orthogonal (with respect to $[x, y]$) subspaces of $\mathbb{R}^6$, we have $\mathbb{R}^6 = \mathcal{M}_+[\dot{+}]\mathcal{M}_-$. It is clear that we cannot enlarge either of the two subspaces. Thus, $\mathcal{M}_+$ should be maximal nonnegative and $\mathcal{M}_-$ should be maximal nonpositive.

When n is even in our example with two Jordan blocks of equal size the invariant maximal nonnegative and maximal nonpositive subspaces are found easily. There is no need to diagonalize some matrix B because of the following: Let $x_1, \ldots, x_{2n}$ and $y_1, \ldots, y_{2n}$ be the two Jordan chains associated with Jordan blocks J_1 and J_2, respectively, both belonging to the same eigenvalue $\lambda = \alpha \pm i\beta$. Let

$$\mathcal{N} = \mathrm{Sp}\{x_1, \ldots, x_n, y_1, \ldots, y_n\}.$$

Then from (3.30) we know that $\langle Hx, y\rangle = 0$ for all $x, y \in \mathcal{N}$, hence $\mathcal{N}$ is H-neutral. Being a neutral subspace it is both a nonnegative and a nonpositive subspace.

3.6 On the structure of H-unitary matrices

Recall that a matrix A is called *H-unitary* if $A^T HA - H = 0$. Obviously the H-unitary matrices form a subclass of the H-expansive matrices. Hence the lemmas in Section 3.2 for the H-expansive case are still applicable to the H-unitary case. However, in the unitary case, it is possible to obtain a further simplification of the matrix H. For instance, it is possible to calculate all the entries in the matrix H, which were previously only denoted by an asterisk (*).

The class of H-unitary matrices has been studied extensively, and canonical forms for pairs (A, H), where $H = H^*$ and A is H-unitary exist both for the complex case (see [24]) and for the real case (see [48] and [66]).

Such canonical forms were developed using the idea of the Cayley transform, and thus usually do not have A in Jordan canonical form. It is the aim of this Section to show that the methods we developed so far for H-expansive matrices may be carried further in the case of H-unitary matrices. We shall not do this all the way through to a canonical form, though we conjecture this is possible as well, leading to a canonical form with A in real or complex Jordan normal form.

We just illustrate the procedure for some examples, hoping to return to the topic in a future publication.

Let us first recall some general properties of H-unitary matrices. If A is H-unitary and $\lambda \in \sigma(A)$ then $(\overline{\lambda})^{-1}$ is an eigenvalue of A as well. Since we are considering the real case, eigenvalues come in four different kinds:

1. Quadruples $\lambda, \overline{\lambda}, \frac{1}{\lambda}, (\overline{\lambda})^{-1}$ with $\lambda \notin \mathbb{R}$ and $|\lambda| \neq 1$.
2. Tuples $\lambda, \frac{1}{\lambda}$ with $\lambda \in \mathbb{R}, \lambda \neq \pm 1$.
3. Tuples $\lambda = \alpha \pm i\beta$, $\alpha^2 + \beta^2 = 1$, $\beta \neq 0$.
4. $\lambda = 1$ or $\lambda = -1$.

Let us agree to denote the (sum of the) root spaces corresponding to such a quadruple, resp. tuple, resp. eigenvalue by $\mathcal{R}_\lambda(A)$. Then these spaces are mutually H-orthogonal (see e.g. [24] or [66]).

Thus we may restrict attention to unitary matrices for which the spectrum is one of these four kinds. In what follows we just focus on $\sigma(A) = \{\alpha \pm i\beta\}$ with $\alpha^2 + \beta^2 = 1$, $\beta \neq 0$.

General remarks

The following general relations can be deduced from the preceding algorithm involving applications of Lemma 3.2.6, Lemma 3.2.7, Lemma 3.2.8 and Lemma 3.2.9 for all entries below and including the main cross diagonal of $(A^T H A - H)$. For A in real Jordan form as in Section 3.3.1 (with one Jordan block) and $H = H^T$, it holds that

$$H_{i,j} = a_{ij} I + b_{ij} H_0 \tag{3.41}$$

$$H_{i,j+1} = \Psi^T H_{i+1,j} - \gamma^T H_{i,j} \tag{3.42}$$

and

$$H_{i,i} = c_i I, \tag{3.43}$$

where c_i has the same meaning as before, namely $c_i = (H_{\lfloor \frac{n+1}{2} \rfloor, \lfloor \frac{n+1}{2} \rfloor})_{ii}$.

We will now show by induction that these relations hold *for all* $i, j = 1, \ldots, n$ (n being odd or even) in the H-unitary case: Relation (3.43) follows from (3.41) and the fact that H must be symmetric. For $i = j = 1$ it immediately follows from Lemma 3.2.5. Assume (3.41) holds for $i+j \leq k$ (and $k < n$). We now show that (3.41) also holds for $i+j \leq k+1$. Recall from (3.21) that

$$(A^T H A - H)_{i,j+1} = H_{i-1,j} + H_{i-1,j+1}\gamma + \gamma^T H_{i,j} + \gamma^T H_{i,j+1}\gamma - H_{i,j+1}.$$

From (3.41) we know that $H_{i-1,j}, H_{i-1,j+1}$ and $H_{i,j}$ are of the form $aI + bH_0$, hence we have from the proof of Lemma 3.2.9 that

$$H_{i-1,j} + H_{i-1,j+1}\gamma + \gamma^T H_{i,j} = 0$$

and

$$\gamma^T H_{i,j+1}\gamma - H_{i,j+1} = 0.$$

Thus, from Lemma 3.2.6 it follows that

$$H_{i,j+1} = a_{i,j+1}I + b_{i,j+1}H_0.$$

Now consider (3.42). This relation holds for all blocks, but it is only of interest for the nonzero block entries. Assume it holds for $i + j \leq k$. Again by (3.21) we have

$$(A^T HA - H)_{i+1,j+1} = H_{i,j} + H_{i,j+1}\gamma + \gamma^T H_{i+1,j} + \gamma^T H_{i+1,j+1}\gamma - H_{i+1,j+1}.$$

Since (3.41) holds, we immediately have by Lemma 3.2.9 that

$$H_{i,j+1} = \Psi^T H_{i+1,j} - \gamma^T H_{i,j}.$$

Hence (3.42) holds for $i + j \leq k + 1$.

The blocks $H_{i,i+1}$ have a particular structure because of (3.42). Indeed, for those blocks we have $H_{i,i+1} = \Psi^T H_{i+1,i} - \gamma^T H_{i,i}$. Using (3.43) and the fact that $H_{i+1,i} = H_{i,i+1}^T$ we see that $H_{i,i+1}$ satisfies

$$H_{i,i+1} = \Psi^T H_{i,i+1}^T - \gamma^T c_i.$$

Lemma 3.6.1 *If $H = aI + bH_0$ satisfies $H = \Psi^T H^T - \gamma^T c$ for some $c \in \mathbb{R}$ then $b = \frac{c}{2\beta} + \frac{\alpha}{\beta}a$. In particular, in that case*

$$H = aI + (\frac{c}{2\beta} + \frac{\alpha}{\beta}a)H_0 = (\frac{1}{2}c + \frac{\alpha}{\beta}\gamma^T)H_0$$

Proof. Insert $H = aI + bH_0$ in $H = \Psi^T H^T - \gamma^T c$, using $\Psi^T = -(\gamma^T)^2 = (\beta^2 - \alpha^2)I + 2\alpha\beta H_0$, $\gamma^T = \alpha I - \beta H_0$, to see that

$$\begin{aligned} aI + bH_0 &= ((\beta^2 - \alpha^2)I + 2\alpha\beta H_0)(aI - bH_0) - (\alpha I - \beta H_0)c \\ &= ((\beta^2 - \alpha^2)a + 2\alpha\beta b - \alpha c)I + ((\beta^2 - \alpha^2)b + 2\alpha\beta a + \beta c)H_0, \end{aligned}$$

where we also used $H_0^2 = -I$. It follows that

$$\begin{aligned} a &= (\beta^2 - \alpha^2)a + 2\alpha\beta b - \alpha c \\ b &= (\beta^2 - \alpha^2)b + 2\alpha\beta a + \beta c. \end{aligned}$$

Now recall $\alpha^2 + \beta^2 = 1$, so

$$(\alpha^2+\beta^2)a = (\beta^2-\alpha^2)a + 2\alpha\beta b - \alpha c$$
$$(\alpha^2+\beta^2)b = (\beta^2-\alpha^2)b + 2\alpha\beta a + \beta c.$$

This is equivalent to

$$2\alpha^2 a - 2\alpha\beta b = -\alpha c$$
$$2\beta^2 a - 2\alpha\beta a = -\beta c.$$

Dividing the first equation by α and the second by β, we see $2\alpha a - 2\beta b = -c$ from both equations. This gives $b = \frac{c}{2\beta} + \frac{\alpha}{\beta}a$ as claimed.

Next, the first form of H now follows directly. To see the second, consider $\beta H H_0 = a\beta H_0 - (\frac{1}{2}c + a\alpha) = -a(\alpha I - \beta H_0) - \frac{1}{2}c = -a\gamma^T - \frac{1}{2}$. Thus,

$$H = \frac{\alpha}{\beta}\gamma^T H_0 + \frac{1}{2}cH_0$$

as desired. □

When n is even, we have $H_{\frac{n}{2},\frac{n}{2}+1} = \Psi^T H^T_{\frac{n}{2},\frac{n}{2}+1} - c_{\frac{n}{2}}\gamma^T$. Thus we have, applying Lemma 3.6.1 with $c_{\frac{n}{2}} = 0$ that

$$H_{\frac{n}{2},\frac{n}{2}+1} = \begin{bmatrix} \beta & \alpha \\ -\alpha & \beta \end{bmatrix} d,$$

for some real number d.

Since every $H_{i,j}$ is determined by blocks in the antidiagonal on which it is, and on the antidiagonal above it (by 3.43), we see that H is completely determined by n real parameters, one for each non-zero antidiagonal.

We now illustrate the unitary case by means of two examples. We will consider a 5×5 block matrix and a 4×4 block matrix, respectively.

Example 3.6.2

Let A be a 5×5 H-unitary matrix, that is $A^THA - H = 0$. Put

$$A = \begin{bmatrix} \gamma & I & & & \\ & \gamma & I & & \\ & & \gamma & I & \\ & & & \gamma & I \\ & & & & \gamma \end{bmatrix}, \quad H = [H_{i,j}]_{i,j=1}^5.$$

Recall that $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, the matrix H is real and $H^T = H$.

From the algorithm in the H-expansive case (refer to Section 3.3) we have the following:

$$
\begin{array}{l}
H_{1,j} = 0 \quad \text{for } j = 1, \ldots, 4 \quad \text{and} \quad H_{1,5} = a_{15}I + b_{15}H_0 \\
H_{2,j} = 0 \quad \text{for } j = 1, \ldots, 3 \quad \text{and} \quad H_{2,4} = a_{24}I + b_{24}H_0 \\
H_{1,5} = \Psi^T H_{2,4} \quad \text{and} \quad H_{2,5} = a_{25}I + b_{25}H_0 \\
H_{3,j} = 0 \quad \text{for } j = 1, 2 \quad \text{and} \quad H_{3,3} = c_3 I \\
H_{2,4} = c_3\Psi^T, \quad H_{3,4} = a_{34}I + b_{34}H_0 \quad \text{and} \quad H_{1,5} = c_3(\Psi^T)^2 \\
H_{4,1} = 0, \quad H_{4,2} = H_{2,4}^T = c_3\Psi, \quad H_{4,3} = H_{3,4}^T \quad \text{and} \quad H_{1,5} = H_{5,1}^T = c_3\Psi^2.
\end{array}
$$

This was how far we could go in the H-expansive case. However, in the unitary case we can now proceed by using Lemma 3.2.9 on the remaining entries. The first entry to consider is:

$$(A^T HA - H)_{3,5} = H_{2,4} + H_{2,5}\gamma + \gamma^T H_{3,4} + \gamma^T H_{3,5}\gamma - H_{3,5} = 0.$$

Applying Lemma 3.2.9 yields

$$H_{2,5} = \Psi^T H_{3,4} - \gamma^T H_{2,4} \quad \text{and} \quad H_{3,5} = a_{35}I + b_{35}H_0.$$

Looking at the $(4,4)$-block entry we obtain

$$(A^T HA - H)_{4,4} = H_{3,3} + H_{3,4}\gamma + \gamma^T H_{4,3} + \gamma^T H_{4,4}\gamma - H_{4,4} = 0.$$

From Lemma 3.2.9 we have

$$H_{3,4} = \Psi^T H_{4,3} - c_3\gamma^T \quad \text{and} \quad H_{4,4} = c_4 I.$$

It follows from equation (3.21) and the application of Lemma 3.2.9 on the $(4,5)$-block entry that

$$H_{3,5} = \Psi^T c_4 - \gamma^T H_{3,4} \quad \text{and} \quad H_{4,5} = a_{45}I + b_{45}H_0.$$

Since $H = H^T$ we know that

$$
\begin{array}{rcl}
H_{5,1} & = & H_{1,5}^T = c_3\Psi^2 \\
H_{5,2} & = & H_{2,5}^T = H_{3,4}^T\Psi - c_3\Psi\gamma \\
H_{5,3} & = & H_{3,5}^T = c_4\Psi - H_{3,4}^T\gamma.
\end{array}
$$

We again use equation (3.21) and Lemma 3.2.9 to find the (4,5)-block entry and hence also the (5,4)-block entry. That is,

$$H_{4,5} = \Psi^T H_{5,4} - c_4\gamma^T \quad \text{and} \quad H_{5,5} = c_5 I.$$

Hence, for $n = 5$, the matrix H reduces to the following form

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & c_3(\Psi^T)^2 \\
0 & 0 & 0 & c_3\Psi^T & \Psi^T H_{3,4} - c_3\gamma^T\Psi^T \\
0 & 0 & c_3 I & \Psi^T H_{3,4}^T - c_3\gamma^T & c_4\Psi^T - \gamma^T H_{3,4} \\
0 & c_3\Psi & H_{3,4}\Psi - c_3\gamma & c_4 I & \Psi^T H_{4,5}^T - c_4\gamma^T \\
c_3\Psi^2 & H_{3,4}^T\Psi - c_3\Psi\gamma & c_4\Psi - H_{3,4}^T\gamma & H_{4,5}\Psi - c_4\gamma & c_5 I
\end{bmatrix}.
$$

Example 3.6.3

Let A and H be real 4×4 block matrices such that $A^T H A - H = 0$. Put,

$$A = \begin{bmatrix} \gamma & I & & \\ & \gamma & I & \\ & & \gamma & I \\ & & & \gamma \end{bmatrix}, \quad H = [H_{i,j}]_{i,j=1}^4,$$

where $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ and $H = H^T$.

As in the previous example, it is possible to compute the remaining entries of the matrix H by using the lemmas in the H-expansive case, together with Lemma 3.2.9. Hence, the corresponding matrix H in this case simplifies to

$$\begin{bmatrix} 0 & 0 & 0 & \Psi^T \mathcal{H} \\ 0 & 0 & \mathcal{H} & c_3 \Psi^T - \gamma^T \mathcal{H} \\ 0 & \mathcal{H}^T & c_3 I & H_{3,4} \\ \mathcal{H}^T \Psi & c_3 \Psi - \mathcal{H}^T \gamma & H_{3,4}^T & c_4 I \end{bmatrix},$$

where $\mathcal{H} = \begin{bmatrix} \beta & \alpha \\ -\alpha & \beta \end{bmatrix} d$, for some real, non-zero d.

Up to this point we have considered the form that H has for any real Jordan basis for A. However, by choosing an appropriate Jordan basis it is possible to achieve a canonical form for H.

First we give a more general form for the new Jordan basis to be constructed. Again we consider the case of one Jordan block $A = J_n(\gamma)$ first, where $\gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ with $\alpha^2 + \beta^2 = 1$, $\beta \neq 0$. Let $x_1, x_2, \ldots, x_{2n-1}, x_{2n}$ be the basis. We denote by $X_1, X_2, \ldots, X_n$ the $2n \times 2$ matrices with columns x_1, x_2; x_3, x_4; $\ldots$; x_{2n-1}, x_{2n}, respectively. Note that

$$AX_1 = X_1\gamma, \ AX_j = X_j\gamma + X_{j-1} \text{ for } j = 2, \ldots, n.$$

Construct a new Jordan basis as follows: Let

$$Z_1 = X_1, \ Z_2 = X_2 + X_1 h_1, \ Z_3 = X_3 + X_2 h_1 + X_1 h_2, \text{ etc,}$$

where the h_j's are 2×2 matrices of the form $aI + bH_0$. So, in general

$$Z_j = X_j + \sum_{i=1}^{j-1} X_i h_{j-i}.$$

From the specific form of h_j we have that $\gamma h_j = h_j \gamma$ (see Lemma 3.2.4). Thus indeed for $j = 2, \ldots, n$ we have

$$\begin{aligned} AZ_j &= AX_j + \sum_{i=1}^{j-1} AX_i h_{j-i} \\ &= X_j\gamma + X_{j-1} + \sum_{i=2}^{j-1}(X_i\gamma + X_{i-1})h_{j-i} + X_1\gamma h_{j-1} \\ &= X_j\gamma + X_{j-1} + \sum_{i=1}^{j-1} X_i h_{j-i}\gamma + \sum_{i=2}^{j-1} X_{i-1}h_{j-i} \\ &= (X_j + \sum_{i=1}^{j-1} X_i h_{j-i})\gamma + (X_{j-1} + \sum_{i=1}^{j-2} X_i h_{j-1-i}) \\ &= Z_j\gamma + Z_{j-1}. \end{aligned}$$

With respect to this new Jordan basis we have a new form for H. Indeed, let $S = [Z_1\, Z_2\, \ldots\, Z_n]$ be the $2n \times 2n$ matrix formed with the Z_i's as block columns. Then $S^{-1}AS = A$ and we compute $S^THS = [(S^THS)_{i,j}]_{i,j=1}^n$. The i,j^{th} entry then is

$$\begin{aligned} (S^THS)_{i,j} &= Z_i^T H Z_j \\ &= (X_i + \sum_{k=1}^{i-1} X_k h_{i-k})^T H (X_j + \sum_{l=1}^{j-1} X_l h_{j-l}) \\ &= H_{i,j} + \sum_{l=1}^{j-1} H_{i,l}h_{j-l} + \sum_{k=1}^{i-1} h_{i-k}^T H_{k,j} + \sum_{k=1}^{i-1}\sum_{l=1}^{j-1} h_{i-k}^T H_{k,l} h_{j-l}. \end{aligned} \quad (3.44)$$

Using these relations it may be possible to choose the matrices h_j, $j = 1, \ldots, n-1$ such that S^THS is brought into a canonical form. We illustrate the idea on two examples, the cases $n = 5$ and $n = 4$ which were already discussed in the previous subsection.

Example 3.6.4

We continue the analysis of Example 3.6.2, so $n = 5$. First take $i = \frac{n+1}{2} = 3$ and $j = \frac{n+3}{2} = 4$ in (3.44). The i,j entry is then the entry just next to the central entry, so in (3.44) the second term becomes $H_{\frac{n+1}{2},\frac{n+1}{2}}h_1 = c_3h_1$. In the third term we have

$$h_1^T H_{\frac{n-1}{2},\frac{n+3}{2}} + h_2^T H_{\frac{n-3}{2},\frac{n+3}{2}} + \ldots = h_1^T H_{\frac{n-1}{2},\frac{n+3}{2}} = h_1^T c_3 \Psi^T.$$

The final double summation in (3.44) becomes zero, so in this case (3.44) becomes

$$H_{3,4} + c_3h_1 + c_3h_1^T\Psi^T = H_{3,4} + c_3(h_1 + h_1^T\Psi^T).$$

Since $H_{3,4}$ is in $\mathcal{E}$ and $h_1 \in \mathcal{E}$ it can be seen by direct computation that h_1 may be chosen such that $(S^THS)_{3,4} = 0$ (provided that $c_3 \neq 0$). Obviously, then

$$(S^THS)_{4,3} = 0 \text{ and } (S^THS)_{5,2} = -c_3\Psi\gamma, \text{ while } (S^THS)_{2,5} = -c_3\gamma^T\Psi^T.$$

In addition, $(S^THS)_{3,5} = c_4\Psi^T$ and $(S^THS)_{5,3} = c_4\Psi$.

Next, we take in (3.44) $i = \frac{n+3}{2} = 4$ and $j = \frac{n+3}{2} = 4$. We obtain from (3.44), using that we have chosen h_1 so that $H_{4,3} = 0$:

$$\begin{aligned}
(S^T HS)_{4,4} &= H_{4,4} + \sum_{l=1}^{3} H_{4,l}h_{4-l} + \sum_{k=1}^{3} h_{4-k}^T H_{k,4} + \sum_{k=1}^{3}\sum_{l=1}^{3} h_{4-k}^T H_{k,l}h_{4-l} \\
&= H_{4,4} + c_3\Psi h_2 + h_2^T c_3\Psi^T + \sum_{l=1}^{3} h_3^T H_{1,l}h_{4-l} \\
&+ \sum_{l=1}^{3} h_2^T H_{2,l}h_{4-l} + \sum_{l=1}^{3} h_1^T H_{3,l}h_{4-l} \\
&= H_{4,4} + c_3(\Psi h_2 + h_2^T\Psi^T + h_1^T h_1) \\
&= c_4 I + c_3(\Psi h_2 + h_2^T\Psi^T + h_1^T h_1).
\end{aligned}$$

Now if $h_1 = a_1 I + b_1 H_0$, then $h_1^T h_1 = (a_1^2 + b_1^2)I$ by Lemma 3.2.4. Setting $h_2 = a_2 I + b_2 H_0$ and recall that $\Psi = -\gamma^2$ we obtain for

$$\Psi h_2 + h_2^T\Psi^T = 2((\beta^2 - \alpha^2)a_2 + 2\alpha\beta b_2)I.$$

Clearly, provided that $c_3 \neq 0$, we can choose a_2 and b_2 so that $(S^T HS)_{4,4} = 0$. This then implies that $(S^T HS)_{5,3} = 0$ and $(S^T HS)_{3,5} = 0$ as we already have $(S^T HS)_{3,4} = 0$. Also, observe that there is still one degree of freedom in choosing a_2 and b_2 if $\alpha \neq 0$.

Further, we take $i = \frac{n+3}{2} = 4$ and $j = \frac{n+5}{2} = 5$ in (3.44). We use the previous steps, so that we may take $H_{4,3} = 0$ and $H_{4,4} = 0$ as well as $H_{5,3} = H_{3,5} = 0$ and $H_{2,5} = -c_3\gamma^T\Psi^T$ and $H_{5,2} = -c_3\Psi\gamma$. With all this we obtain in (3.44):

$$\begin{aligned}
(S^T HS)_{4,5} &= H_{4,5} + \sum_{l=1}^{4} H_{4,l}h_{5-l} + \sum_{k=1}^{3} h_{4-k}^T H_{k,5} + \sum_{k=1}^{3}\sum_{l=1}^{4} h_{4-k}^T H_{k,l}h_{5-l} \\
&= H_{4,5} + c_3\Psi h_3 + h_3^T c_3(\Psi^T)^2 - h_2^T c_3\gamma^T\Psi^T + \sum_{l=1}^{4} h_3^T H_{1,l}h_{5-l} \\
&+ \sum_{l=1}^{4} h_2^T H_{2,l}h_{5-l} + \sum_{l=1}^{4} h_1^T H_{3,l}h_{5-l} \\
&= H_{4,5} + c_3(\Psi h_3 + h_3^T(\Psi^T)^2) - h_2^T c_3\gamma^T\Psi^T + h_2^T c_3\Psi^T h_1 + c_3 h_1^T h_2.
\end{aligned}$$

Again viewing this as an equation in h_3 we see that we can solve this to be zero as well. So we may take $H_{4,5} = 0$, and consequently $H_{5,4} = 0$.

Lastly, consider $i = 5, j = 5$ in (3.44) using the previous steps:

$$\begin{aligned}(S^THS)_{5,5} &= H_{5,5} + \sum_{l=1}^{4} H_{5,l}h_{5-l} + \sum_{k=1}^{4} h_{5-k}^T H_{k,5} + \sum_{k=1}^{4}\sum_{l=1}^{4} h_{5-k}^T H_{k,l} h_{5-l} \\ &= H_{5,5} + c_3\Psi^2 h_4 - c_3\Psi\gamma h_3 + h_4^T c_3 (\Psi^T)^2 - c_3 h_3^T \gamma^T \Psi^T \\ &+ \sum_{l=1}^{4} h_4^T H_{1,l} h_{5-l} + \sum_{l=1}^{4} h_3^T H_{2,l} h_{5-l} + \sum_{l=1}^{4} h_2^T H_{3,l} h_{5-l} + \sum_{l=1}^{4} h_1^T H_{4,l} h_{5-l} \\ &= H_{5,5} + c_3(\Psi^2 h_4 + h_4^T (\Psi^T)^2) - c_3(h_3^T\gamma^T\Psi^T + \Psi\gamma h_3) \\ &+ c_3(h_3^T\Psi^T h_1 + h_1^T \Psi h_3) + c_3 h_2^T h_2.\end{aligned}$$

We see that this is a multiple of the identity, and moreover, it is linear in h_4, and we can find h_4 so that this is zero.

Finally, by considering a constant multiple of the Jordan chains, we may scale c_3 to either $+1$ or -1. We arrive at the following form for the pair (A, H), where

$$A = \begin{bmatrix} \gamma & I & & & \\ & \gamma & I & & \\ & & \gamma & I & \\ & & & \gamma & I \\ & & & & \gamma \end{bmatrix}, \quad \gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

and H is then

$$H = \epsilon \begin{bmatrix} 0 & 0 & 0 & 0 & (\Psi^T)^2 \\ 0 & 0 & 0 & \Psi^T & (\gamma^T)^3 \\ 0 & 0 & I & 0 & 0 \\ 0 & \Psi & 0 & 0 & 0 \\ \Psi^2 & \gamma^3 & 0 & 0 & 0 \end{bmatrix},$$

with $\Psi = -\gamma^2$ and $\epsilon = \pm 1$.

In a very similar way, we can reduce the simple form for the pair (A, H) of Example 3.6.3, that is the 4×4 case, to the following, where

$$A = \begin{bmatrix} \gamma & I & & \\ & \gamma & I & \\ & & \gamma & I \\ & & & \gamma \end{bmatrix}, \quad \gamma = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

and the H in this case will be

$$H = \epsilon \begin{bmatrix} 0 & 0 & 0 & \Psi^T\mathcal{H} \\ 0 & 0 & \mathcal{H} & -\gamma^T\mathcal{H} \\ 0 & \mathcal{H}^T & 0 & 0 \\ \mathcal{H}^T\Psi & -\mathcal{H}^T\gamma & 0 & 0 \end{bmatrix},$$

with $\begin{bmatrix} \beta & \alpha \\ -\alpha & \beta \end{bmatrix}$ and $\epsilon = \pm 1$.

Simple form for H in the general case

The next step in the unitary case is to see what is the most simple form for the matrix H when A consists of more than one Jordan block. For this we need to compute all the entries in the blocks below the main diagonal. As an example we will first look at the case when the matrix A consists of two real Jordan blocks. Let

$$A = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix},$$

where

$$J_p = \begin{bmatrix} \gamma & I & & & \\ & \gamma & I & & \\ & & \ddots & & \\ & & & \gamma & I \\ & & & & \gamma \end{bmatrix} \quad \text{for} \quad p = 1, 2$$

and γ as before. We will illustrate the example, by letting J_1 be 3×3 and J_2 be 2×2 Jordan block matrices with respect to real Jordan bases $\{x_1, \ldots, x_6\}$ and $\{y_1, \ldots, y_4\}$, respectively. We can apply the appropriate lemmas to the blocks on the main diagonal and find the corresponding form of the matrix H on the diagonal. Since $H = H^T$ we only need to look at the $(2,1)$-block entry of $A^T H A - H$, that is, equation (3.31) of Section 3.2. From the same section we know the entries $\tilde{H}_{1,3}$ and $\tilde{H}_{2,2}$, but we could not find the $\tilde{H}_{2,3}$ entry in the H-expansive case. However, we can now apply Lemma 3.2.8 to this entry. Hence we have the following form for the matrix H:

$$\left[\begin{array}{ccc|cc} 0 & 0 & c_2\Psi^T & 0 & 0 \\ 0 & c_2 I & \Psi^T H_{3,2} - c_2\gamma^T & 0 & \tilde{H}_{2,2}^T \\ c_2\Psi & H_{3,2}^T\Psi - c_2\gamma & c_3 I & \tilde{H}_{2,2}^T\Psi & \tilde{H}_{2,3}^T \\ \hline 0 & 0 & \Psi^T\tilde{H}_{2,2} & 0 & \mathcal{H} \\ 0 & \tilde{H}_{2,2} & \tilde{H}_{2,3} & \mathcal{H}^T & c_2' I \end{array}\right], \tag{3.45}$$

with

$$\begin{aligned} c_2 &= 2(\beta b_{23} - \alpha a_{23}), \quad c_3 = (H_{3,3})_{11} \\ \mathcal{H} &= d(\beta I + \alpha H_0), \quad c_2' = (H_{2,2})'_{11}. \end{aligned}$$

Remember that we now know that all entries in H is of the form $aI + bH_0$ for some a and b.

Knowing that we can find all the entries in the blocks below the main diagonal blocks (hence also those entries above the main diagonal blocks, since $H = H^T$) we will now change the Jordan basis of A such that the entries below the main diagonal blocks become zero. We organize the real Jordan basis $\{x_1, \ldots, x_6, y_1, \ldots, y_4\}$ in the following way: We denote by X_1, X_2, X_3 the $n \times 2$ matrices with columns $x_1, x_2; x_3, x_4; x_5, x_6$, respectively. Likewise Y_1, Y_2 are the $n \times 2$ matrices $Y_1 = [y_1, y_2], Y_2 = [y_3, y_4]$. Note that:

$$\begin{aligned} AX_3 &= X_3\gamma + X_2, \quad AX_2 = X_2\gamma + X_1, \quad AX_1 = X_1\gamma \\ AY_2 &= Y_2\gamma + Y_1, \quad AY_1 = Y_1\gamma \end{aligned}$$

Construct a new Jordan basis as follows: Let

$$\begin{aligned} Z_3 &= X_3 + Y_2h + Y_1k \\ Z_2 &= X_2 + Y_1h \\ Z_1 &= X_1, \end{aligned}$$

where $h, k \in \mathbb{R}^{2\times 2}$ are of the form $aI + bH_0$. From the specific form of h, k and γ we have that $\gamma h = h\gamma$ and $\gamma k = k\gamma$. (See Lemma 3.2.4). We now find that

$$\begin{aligned} AZ_3 &= AX_3 + AY_2h + AY_1k \\ &= X_3\gamma + X_2 + Y_2\gamma h + Y_1h + Y_1\gamma k \\ &= (Z_3 - Y_2h - Y_1k)\gamma + Z_2 - Y_1h + Y_2\gamma h + Y_1h + Y_1\gamma k \\ &= Z_3\gamma + Z_2. \end{aligned}$$

Similarly we obtain that

$$AZ_2 = Z_2\gamma + Z_1 \quad \text{and} \quad AZ_1 = Z_1\gamma.$$

With the new Jordan basis it is now possible to make $\tilde{H}_{2,2} = a_{22}I + b_{22}H_0 = 0$ as well as $\tilde{H}_{2,3} = a_{23}I + b_{23}H_0 = 0$. Then it immediately follows that the $(1,3)$-block entry of the $(2,1)$-block of the matrix in (3.45) is zero, since it is related to the $(2,2)$-block entry. So, let us consider the $(2,2)$-block entry of the $(2,1)$-block in matrix (3.45):

$$\begin{aligned} Y_2^T H Z_2 &= Y_2^T H(X_2 + Y_1h) \\ &= Y_2^T H X_2 + Y_2^T H Y_1 h \\ &= \tilde{H}_{2,2} + \mathcal{H}^T h. \end{aligned}$$

Hence, if we take

$$h = -(\mathcal{H}^T)^{-1}\tilde{H}_{2,2}$$

then $Y_2^T H Z_2 = 0$. For the $(2,3)$-block entry to be zero we have to choose k such that

$$\begin{aligned} 0 = Y_2^T H Z_3 &= Y_2^T H(X_3 + Y_2h + Y_1k) \\ &= Y_2^T H X_3 + Y_2^T H Y_2 h + Y_2^T H Y_1 k \\ &= \tilde{H}_{2,3} + c_2' h + \mathcal{H}^T k. \end{aligned}$$

Hence, recalling that we have $h = -(\mathcal{H}^T)^{-1}\tilde{H}_{2,2}$ then

$$k = c_2'(\mathcal{H}^T)^{-1}\tilde{H}_{2,2}(\mathcal{H}^T)^{-1} - \tilde{H}_{2,3}(\mathcal{H}^T)^{-1}.$$

Therefore by changing the Jordan basis we can find zero entries in the blocks below (and above) the main diagonal blocks, such that our matrix H becomes:

$$\left[\begin{array}{ccc|cc} 0 & 0 & c_2\Psi^T & 0 & 0 \\ 0 & c_2 I & \Psi^T H_{3,2} - c_2\gamma^T & 0 & 0 \\ c_2\Psi & H_{3,2}^T\Psi - c_2\gamma & c_3 I & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \mathcal{H} \\ 0 & 0 & 0 & \mathcal{H}^T & c_2' I \end{array}\right]. \tag{3.46}$$

For each of these two blocks we can now separately apply the procedure sketched in the previous subsection and finally arrive at the simple form for the pair (A, H):

$$A = \left[\begin{array}{ccc|cc} \gamma & I & & & \\ & \gamma & I & & \\ & & \gamma & & \\ \hline & & & \gamma & I \\ & & & & \gamma \end{array}\right] \tag{3.47}$$

and

$$\tilde{H} = \left[\begin{array}{ccc|cc} 0 & 0 & \epsilon_1\Psi^T & 0 & 0 \\ 0 & \epsilon_1 I & -\epsilon_1\gamma^T & 0 & 0 \\ \epsilon_1\Psi & -\epsilon_1\gamma & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \epsilon_2\mathcal{H} \\ 0 & 0 & 0 & \epsilon_2\mathcal{H}^T & 0 \end{array}\right], \tag{3.48}$$

with $\mathcal{H} = \begin{bmatrix} \beta & \alpha \\ -\alpha & \beta \end{bmatrix}$ and $\epsilon_1, \epsilon_2 = \pm 1$.

The procedure outlined in the previous subsections will eventually lead to a form for (A, H), which then has to be shown to be canonical.

It should be observed that the case of the eigenvalues ± 1 is more complicated than what we have considered separately. Indeed, the canonical form in [66] shows that Jordan blocks in the Jordan canonical form with eigenvalue ± 1 and even size must come in coupled pairs. This complicates the procedure.

Chapter 4

Rank-one perturbations of H-positive real matrices

4.1 Introduction

In this chapter we consider perturbations of rank-one of positive real matrices. So, let H be a complex Hermitian invertible $n \times n$ matrix, and let A be H-positive real, that is, $HA + A^*H = D \geq 0$.

A rank-one perturbation of A is given by $B = A + uv^*$, where u and v are in $\mathbb{C}^n$. We shall primarily consider perturbations that have the additional structure that uv^* is itself H-positive real. We are interested in the generic behaviour of eigenvalues under such structured rank-one perturbations. The genericity of the perturbation can be understood in various ways. In this chapter it is understood in the sense of algebraic geometry.

The following formulation of genericity is taken from [53]. We say that a set $W \subseteq \mathbb{F}^n$ (where $\mathbb{F}$ denotes either the field of complex numbers or the field of real numbers) is algebraic if there exists a finite set of polynomials $f_1(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n)$ with coefficients in $\mathbb{F}$ such that a vector $[a_1, \ldots, a_n]^T \in \mathbb{F}^n$ belongs to W if and only if

$$f_j(a_1, \ldots, a_n) = 0, \quad j = 1, 2, \ldots, k.$$

In particular, the empty set is algebraic and $\mathbb{F}^n$ is algebraic. We say that a set $W \subseteq \mathbb{F}^n$ is generic if W is not empty and the complement $\mathbb{F}^n \setminus W$ is contained in an algebraic set which is not $\mathbb{F}^n$. Note here that the union of finitely many algebraic sets is again algebraic. Clearly, if the set $W \subseteq \mathbb{F}^n$ is generic and if $S \in \mathbb{F}^{n \times n}$ is invertible then SW is also generic. In the following, we say that a set $W \subseteq \mathbb{F}^n \times \mathbb{F}^n$ is generic if W, canonically identified with a subset of $\mathbb{F}^{2n}$ is generic as a subset of $\mathbb{F}^{2n}$.

In addition further properties of *generic sets* can be found in Section 2 of [53].

4.2 Preliminaries

As a first important observation we have the following proposition.

Proposition 4.2.1 *A rank-one matrix uv^* is H-positive real if and only if $v = c \cdot Hu$ for a non-zero constant c, with* $\operatorname{Re} c \geq 0$.

Proof. We have $Huv^* + vu^*H \geq 0$ if and only if $x^*(Huv^* + vu^*H)x \geq 0$ for all vectors x. Now $Huv^* + vu^*H$ is a-priori of rank two at most. Moreover, if x is in the kernel of $Huv^* + vu^*H$, then the inner product $x^*(Huv^* + vu^*H)x = 0$. So we only need to consider x in the orthogonal complement of the kernel of $Huv^* + vu^*H$. Since this matrix is Hermitian, we only need to consider x in the image of $Huv^* + vu^*H$, that is, we may consider $x = c_1Hu + c_2v$. Thus, we compute

$$\begin{aligned} & (\overline{c_1}u^*H + \overline{c_2}v^*)(Huv^* + vu^*H)(c_1Hu + c_2v) \\ = \; & |c_1|^2u^*H(Huv^* + vu^*H)Hu + \overline{c_1}c_2u^*H(Huv^* + vu^*H)v \\ + \; & \overline{c_2}c_1v^*(Huv^* + vu^*H)Hu + |c_2|^2v^*(Huv^* + vu^*H)v \end{aligned}$$

This is a quadratic form in c_1, c_2 and the corresponding two by two matrix is

$$\begin{aligned} D &= \begin{bmatrix} u^*H(Huv^* + vu^*H)Hu & u^*H(Huv^* + vu^*H)v \\ v^*(Huv^* + vu^*H)Hu & v^*(Huv^* + vu^*H)v \end{bmatrix} \\ &= \begin{bmatrix} 2\|Hu\|^2\text{Re } v^*Hu & \|Hu\|^2\|v\|^2 + (u^*Hv)^2 \\ \|Hu\|^2\|v\|^2 + (v^*Hu)^2 & 2\|v\|^2\text{Re } v^*Hu \end{bmatrix}. \end{aligned}$$

Thus, we have that $Huv^* + vu^*H \geq 0$ if and only if $D \geq 0$. Now for a two by two matrix to be positive semidefinite, it is necessary and sufficient that the determinants of the principal submatrices are greater than or equal to zero. Concerning D this comes down to just two conditions,

$$\begin{aligned} &\text{Re } v^*Hu \geq 0, \\ &2\|Hu\|^2\|v\|^2(\text{Re } v^*Hu)^2 - \|Hu\|^4\|v\|^4 - |v^*Hu|^4 \geq 0. \end{aligned}$$

From the second condition we obtain

$$-(|v^*Hu|^2 - \|Hu\|^2\|v\|^2)^2 \geq 0.$$

Hence, $|v^*Hu|^2 = \|Hu\|^2\|v\|^2$. It follows that equality holds in the Cauchy-Schwarz inequality, and this is the case if and only if v is a multiple of Hu, i.e., $v = c \cdot Hu$. Then $uv^* = cuu^*H$ for some complex number c. Since $Huv^* + vu^*H \geq 0$ we conclude that Re $c \geq 0$, and since uv^* is of rank-one, we conclude c is non-zero. □

Next, we extend Proposition 4.2.1 to H-positive real *matrices of rank-k*. But first the following lemma is required in the proof of Proposition 4.2.3.

Lemma 4.2.2 *Let $H = H^*$ be an invertible $n \times n$ matrix. For $n \times k$ matrices U and V and invertible matrices S and T of sizes $k \times k$ and $n \times n$ respectively, let*

$$H_1 = T^*HT, \qquad U_1 = T^{-1}US^{-1}, \qquad V_1^* = SV^*T.$$

Then,

$$HUV^* + VU^*H \geq 0 \Leftrightarrow H_1U_1V_1^* + V_1U_1^*H_1 \geq 0.$$

Further, $\operatorname{rank} U = \operatorname{rank} U_1$, $\operatorname{rank} V = \operatorname{rank} V_1$, $\operatorname{rank} U_1V_1^* = \operatorname{rank} UV^*$ *and* $\operatorname{rank}(HUV^* + VU^*H) = \operatorname{rank}(H_1U_1V_1^* + V_1U_1^*H_1)$.

Proof. The positive real statement follows from the identity:

$$\begin{aligned} T^*(HUV^* + VU^*H)T &= T^*H(TT^{-1})U(S^{-1}S)V^*T \\ &+ T^*V(S^*(S^*)^{-1})U^*((T^*)^{-1}T^*)HT \\ &= H_1U_1V_1^* + V_1U_1^*H_1. \end{aligned}$$

The rank statements follow from the fact that the invertibility of S and T preserve rank. □

Proposition 4.2.3 *Let U, V be $n \times k$ matrices such that UV^* has rank-k. Then UV^* is H-positive real if and only if there exists a $k \times k$ invertible matrix E such that $V = HUE$ and $E^* + E \geq 0$. Further,* $\operatorname{rank}(HUV^* + VU^*H) = \operatorname{rank}(E^* + E)$.

Proof. Let $V = HUE$, where E has the properties in the statement of the proposition. Then,

$$\begin{aligned} HUV^* + VU^*H &= HUE^*U^*H + HUEU^*H \\ &= HU(E^* + E)U^*H. \end{aligned}$$

Since, $E^* + E \geq 0$, it follows that $HUV^* + VU^*H \geq 0$.

Conversely, assume that UV^* is H-positive real. From $\operatorname{rank}(UV^*) = k$, it follows that $\operatorname{rank}(U) = \operatorname{rank}(V^*) = k$. Therefore we can find invertible matrices S and T such that

$$SV^*T = \begin{bmatrix} I_k & 0 \end{bmatrix}.$$

Using the matrices S and T as in Lemma 4.2.2, we may assume without loss of generality that $V^* = \begin{bmatrix} I_k & 0 \end{bmatrix}$ and that our matrices H, U and V were obtained similarly to the matrices H_1, U_1 and V_1 in Lemma 4.2.2. If we let $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$, where U_1 is a $k \times k$ block and U_2 is a $(n-k) \times k$ block, then

$$UV^* = \begin{bmatrix} U_1 & 0 \\ U_2 & 0 \end{bmatrix} \quad \text{and} \quad VU^* = \begin{bmatrix} U_1^* & U_2^* \\ 0 & 0 \end{bmatrix}.$$

If we let $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$, where H_{11} is a $k \times k$ block, then

$$HUV^* + VU^*H = \begin{bmatrix} D + D^* & U_1^*H_{12} + U_2^*H_{22} \\ H_{21}U_1 + H_{22}U_2 & 0 \end{bmatrix}, \tag{4.1}$$

where we let $D = H_{11}U_1 + H_{12}U_2$. Since we assume that $HUV^* + VU^*H \geq 0$, it follows that $H_{21}U_1 + H_{22}U_2 = U_1^*H_{12} + U_2^*H_{22} = 0$. Thus we have

$$HUV^* + VU^*H = \begin{bmatrix} D + D^* & 0 \\ 0 & 0 \end{bmatrix}.$$

This implies that $D + D^* \geq 0$ and that $\operatorname{rank}(HUV^* + VU^*H) = \operatorname{rank}(D + D^*)$, where $D + D^*$ is a $k \times k$ matrix. Also, because

$$HU = \begin{bmatrix} D \\ 0 \end{bmatrix},$$

it is clear that $\text{rank}(D) = \text{rank}(HU) = \text{rank}(U) = k$. Since D is a $k \times k$ matrix of rank-k, it should be invertible. Let $E = D^{-1}$. An easy calculation shows that

$$\langle (E^* + E)x, x\rangle = \langle (D + D^*)(D^{-1}x), D^{-1}x\rangle \geq 0$$

for all x. Thus $E^* + E \geq 0$. Further,

$$HUE = \begin{bmatrix} D \\ 0 \end{bmatrix} D^{-1} = \begin{bmatrix} I_k \\ 0 \end{bmatrix} = V$$

and

$$E + E^* = D^{-1}(D^* + D)(D^*)^{-1}.$$

It now follows that

$$\text{rank}(E + E^*) = \text{rank}(D^* + D) = \text{rank}(HUV^* + VU^*H).$$

□

Remark 4.2.4 Our result of Proposition 4.2.1 also follows from Proposition 4.2.3: Assume $\text{rank}(uv^*) = 1$. Thus in Proposition 4.2.3 we can replace the $n \times k$ matrices by the $n \times 1$ vectors. By Proposition 4.2.3 the matrix uv^* is H-positive real if and only if there exists an invertible 1×1 matrix E (i.e. a number α) such that $\overline{\alpha} + \alpha \geq 0$, i.e., $\alpha \geq 0$, such that $v = \alpha \cdot Hu$. Since uv^* is a non-zero matrix, we have $\alpha \neq 0$.

Although the above results are proved for the complex case, it also holds in the real case, that is, when H is a real symmetric invertible $n \times n$ matrix and A being H positive real, i.e. $HA + A^T H \geq 0$.

4.3 Main result

In this section we prove one of the main results of this chapter. In the sequel we shall restrict our attention to rank-one perturbation of the type tuu^*H with *real* $t > 0$. This is in particular motivated by the real case, which we shall discuss later in this chapter. In order to state this, we need to consider the Jordan form of the H-positive real matrix A. We denote by λ_j, $j = 1, \ldots, p$ the eigenvalues of A, and for every j, we denote by k_j the geometric multiplicity of λ_j, and by $n_{j,1} \geq n_{j,2} \geq \cdots \geq n_{j,k_j}$ the partial multiplicities of A corresponding to λ_j.

Theorem 4.3.1 *Let A be H-positive real. Then, for generic u and H-positive real rank-one perturbation $B(t) = A + tuu^*H$ $(t > 0)$, we have the following:*

1. *For each $j = 1, \ldots, p$, and for every positive t, $B(t)$ has partial multiplicities $n_{j,2} \geq \cdots \geq n_{j,k_j}$ corresponding to the eigenvalues λ_j.*

2. *For all but finitely many $t > 0$, $B(t)$ has $\sum_{j=1}^{p} n_{j,1}$ simple eigenvalues which are not eigenvalues of A.*

   For all u it follows that:

3. *For all $t > 0$, all the pure imaginary eigenvalues of $B(t)$ are eigenvalues of A. Thus the simple eigenvalues of $B(t)$ (for all but possibly a finite number of $t > 0$) mentioned in Part 2 are not pure imaginary.*

4. *The eigenvectors of $B(t)$ corresponding to an eigenvalue λ which is not an eigenvalue of A are strictly H-positive if $\mathrm{Re}\ \lambda > 0$ and strictly H-negative if $\mathrm{Re}\ \lambda < 0$, that is if $Ax = \lambda x$ with $x \neq 0$ and $\mathrm{Re}\ \lambda > 0$, then $\langle Hx, x\rangle > 0$, and if $\mathrm{Re}\ \lambda < 0$, then $\langle Hx, x\rangle < 0$.*

We first observe that the major part of the theorem is in line with a result concerning unstructured rank-one perturbations, see [68], also [14]. In fact, the only statement in the theorem that is different from the statements in these references is the statement concerning the location of the simple eigenvalues, i.e., Parts 3 and 4.

Example 4.3.2

As a first example of the result of the theorem, consider

$$A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \qquad H = I_2$$

Then, according to Proposition 4.2.1, a rank-one H-positive real matrix is of the form uu^*, so the *perturbed matrix* becomes

$$B = A + uu^* = \begin{bmatrix} |u_1|^2 & 2 + u_1\overline{u}_2 \\ -2 + \overline{u}_1 u_2 & |u_2|^2 \end{bmatrix}.$$

If this matrix would have pure-imaginary eigenvalues, then its trace would have to be zero, i.e., $|u_1|^2 + |u_2|^2 = 0$. This is only possible if $u_1 = u_2 = 0$. So, we see that for this example no rank-one positive real perturbation can have pure-imaginary eigenvalues, as predicted by the theorem.

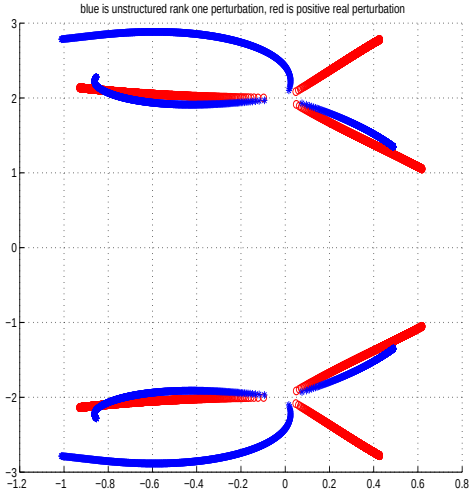
Observe that the result is not true if we replace the H-positive real rank-one perturbation by a general rank-one perturbation, that is, for generic u and v in $\mathbb{R}^n$ and $C(t) = A + tuv^*$ it is very well possible that $C(t)$ has pure imaginary eigenvalues for some positive values of t. This is illustrated in the figures below, where the graphs are plotted in the complex plane and the paths of the perturbed eigenvalues are shown as a function of time. The first figure presents the case where

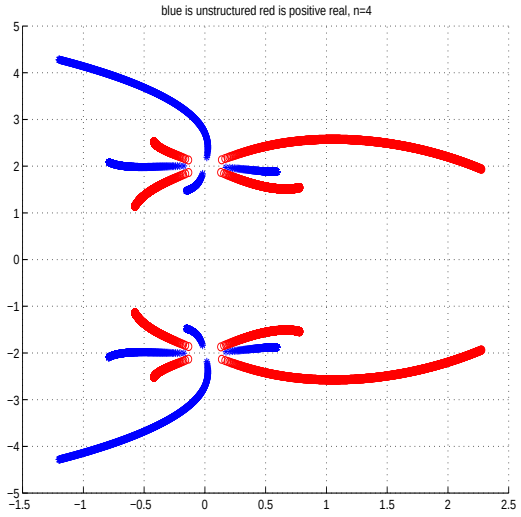
$$A = \begin{bmatrix} \gamma & I & 0 \\ 0 & \gamma & I \\ 0 & 0 & \gamma \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & 0 & -I \\ 0 & I & 0 \\ -I & 0 & 0 \end{bmatrix}, \tag{4.2}$$

while the second figure presents the case where

$$A = \begin{bmatrix} \gamma & I & 0 & 0 \\ 0 & \gamma & I & 0 \\ 0 & 0 & \gamma & I \\ 0 & 0 & 0 & \gamma \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & 0 & 0 & -I \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ -I & 0 & 0 & 0 \end{bmatrix}. \tag{4.3}$$

In both cases $\gamma = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$.





The proofs of Part 1 and 2 in our context follows later. We start with the proofs of Parts 3 and 4 of Theorem 4.3.1.

Proof. Part 3: Suppose that

$$(A + tuu^*H)x = \lambda_0 x, \qquad \operatorname{Re}\lambda_0 = 0,$$

for some non-zero complex vector $x \in \mathbb{C}^n$, and some $t > 0$. Multiply the latter equation with H from the left, and take the inner product with x to obtain

$$\langle (HA + tHuu^*H)x, x\rangle = \lambda_0 \langle Hx, x\rangle.$$

The right hand side has real part zero, since $\langle Hx, x\rangle$ is a real number and λ_0 is pure-imaginary. The real part of the left hand side is given by

$$\begin{aligned} & \operatorname{Re}\langle HAx, x\rangle + t\langle Huu^*Hx, x\rangle = \\ = \ & \frac{1}{2}(\langle HAx, x\rangle + \langle x, HAx\rangle) + t\langle u^*Hx, u^*Hx\rangle = \\ = \ & \frac{1}{2}(\langle HAx, x\rangle + \langle A^*Hx, x\rangle) + t|u^*Hx|^2 = \\ = \ & \frac{1}{2}\langle Dx, x\rangle + t|u^*Hx|^2, \end{aligned}$$

where $D = HA + A^*H$. Now since $D \geq 0$, we see that the latter expression can only be zero if both terms are zero. Thus, in particular, $u^*Hx = 0$. But then it follows from $(A + tuu^*H)x = \lambda_0 x$ that in fact $Ax = \lambda_0 x$. That is, the pure-imaginary eigenvalues of $A + tuu^*H$ for any u and any $t > 0$ are those of A.

Part 4: The computation above actually can be used to prove Part 4 as well. Let λ be an eigenvalue of $B(t)$ and suppose λ is not an eigenvalue of A. Following the arguments in the proof of Part 3, we see that

$$\operatorname{Re}\lambda\langle Hx, x\rangle = \frac{1}{2}\langle Dx, x\rangle + t|u^*Hx|^2 \geq 0, \tag{4.4}$$

where $D = HA + A^*H$, $t > 0$ and $x \neq 0$ is an eigenvector corresponding to λ. Since we assume that λ is not an eigenvalue of A, it follows from Part 3 that if $\operatorname{Re}\lambda > 0$ (i.e. λ is in the open right half plane), then we have $\langle Hx, x\rangle \geq 0$. Also, as in the proof of Part 3 we have seen that $\langle Hx, x\rangle > 0$ or else $u^*Hx = 0$ and then λ is an eigenvalue of A. Similarly, $\langle Hx, x\rangle < 0$ if $\operatorname{Re}\lambda < 0$.

Thus we see that the eigenvalues of $B(t)$ which are not eigenvalues of A, as mentioned in Part 3 of the theorem have the following property: all eigenvectors corresponding to those eigenvalues that are in the open left half plane are strictly H-negative vectors, all eigenvectors corresponding to those eigenvalues that are in the open right half plane are strictly H-positive vectors. □

Next, we consider perturbations of rank larger than one. We shall restrict our attention to perturbations of the type tUE^*U^*H with real $t > 0$. This is in particular motivated by the real case, which we shall discuss later in this chapter.

Theorem 4.3.3 *Let A be H-positive real. Then, for any $n \times k$ matrix U and for any invertible $k \times k$ matrix E with $E + E^* > 0$, the H-positive real rank-k perturbation $B(t) = A + tUE^*U^*H$ $(t > 0)$ has the following properties:*

1. *For all $t > 0$, all the pure imaginary eigenvalues of $B(t)$ are eigenvalues of A.*
2. *The eigenvectors of $B(t)$ corresponding to an eigenvalue λ which is not an eigenvalue of A are strictly H-positive if $\operatorname{Re}\lambda > 0$ and strictly H-negative if $\operatorname{Re}\lambda < 0$.*

Proof. Part 1: Suppose that

$$(A + tUE^*U^*H)x = \lambda_0 x, \qquad \operatorname{Re}\lambda_0 = 0,$$

for some non-zero complex vector $x \in \mathbb{C}^n$, and some $t > 0$. Multiply the latter equation with H from the left, and take the inner product with x to obtain

$$\langle (HA + tHUE^*U^*H)x, x\rangle = \lambda_0 \langle Hx, x\rangle.$$

The right hand side has real part zero, since $\langle Hx, x\rangle$ is a real number and λ_0 is pure-imaginary. The real part of the left hand side is given by

$$\begin{aligned} & \mathrm{Re}\,\langle HAx, x\rangle + t\langle HUE^*U^*Hx, x\rangle \\ = \; & \frac{1}{2}(\langle HAx, x\rangle + \langle x, HAx\rangle) + t\mathrm{Re}\langle E^*U^*Hx, U^*Hx\rangle \\ = \; & \frac{1}{2}\langle (HA + A^*H)x, x\rangle + t\{\langle E^*U^*Hx, U^*Hx\rangle + \langle EU^*Hx, U^*Hx\rangle\} \\ = \; & \frac{1}{2}\langle (HA + A^*H)x, x\rangle + t\langle (E^* + E)U^*Hx, U^*Hx\rangle \\ = \; & \frac{1}{2}\langle (HA + A^*H)x, x\rangle + t\|(E + E^*)^{1/2}U^*Hx\|^2. \end{aligned}$$

Now, since $HA + A^*H \geq 0$, we see that the latter expression can only be zero if both terms are zero. In particular, $(E + E^*)^{1/2}U^*Hx = 0$ and the fact that $E + E^*$ has rank-k implies that $U^*Hx = 0$. But then it follows from $(A + tUE^*U^*H)x = \lambda_0 x$ that in fact $Ax = \lambda_0 x$. That is, the pure-imaginary eigenvalues of $A + tUE^*U^*H$ for any U and any $t > 0$ are those of A.

The proof of Part 2 of this theorem follows similarly as in the proof of Part 4 in Theorem 4.3.1. □

Remark 4.3.4 It might seem reasonable to conjecture that the number of eigenvalues moving into the left/right half plane is determined by the signs corresponding to odd blocks with pure-imaginary eigenvalues in the Jordan canonical form of A. This turns out not to be true, at least not in case $n_{j1} = n_{j2}$ and n_{j1} is odd and there is opposite signs. As an example, consider the case where $\lambda_j = 0$, and $n_{j1} = n_{j2} = 1$, that is:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Then $A + uu^*H = \begin{bmatrix} |u_1|^2 & -u_1\overline{u}_2 \\ u_2\overline{u}_1 & -|u_2|^2 \end{bmatrix}$, which has eigenvalues 0 and $|u_1|^2 - |u_2|^2$. Obviously, the latter number can be positive or negative, depending on u, so there is no dependence on the signs in H.

The next example illustrates Part 3 of Theorem 4.3.1 in a very explicit computation.

Example 4.3.5

Consider first the pair of matrices given in (4.2). We shall denote the vector u by $u^* = \begin{bmatrix} u_1^* & u_2^* & u_3^* \end{bmatrix}$, where u_1, u_2 and u_3 are vectors in $\mathbb{C}^2$, and $\gamma = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$. In this case $\det(\lambda - A) = (\lambda^2 + b^2)^3$ and if $\lambda \neq \pm ib$ then one computes that

$$(\lambda - A)^{-1} = \begin{bmatrix} (\lambda-\gamma)^{-1} & (\lambda-\gamma)^{-2} & (\lambda-\gamma)^{-3} \\ 0 & (\lambda-\gamma)^{-1} & (\lambda-\gamma)^{-2} \\ 0 & 0 & (\lambda-\gamma)^{-1} \end{bmatrix}.$$

One has, using $\det(\lambda - \gamma) = \lambda^2 + b^2$,

$$\begin{aligned} (\lambda-\gamma)^{-1} &= \frac{1}{\lambda^2+b^2}(\lambda+\gamma), \\ (\lambda-\gamma)^{-2} &= \frac{1}{(\lambda^2+b^2)^2}(\lambda^2+2\lambda\gamma+\gamma^2), \\ (\lambda-\gamma)^{-3} &= \frac{1}{(\lambda^2+b^2)^3}(\lambda^3+3\lambda^2\gamma+3\lambda\gamma^2+\gamma^3). \end{aligned}$$

Then the next computation comes in handy, where we use Lemma 4.6.1 in the third step

$$\begin{aligned} & \det(\lambda I - (A + tuu^*H)) \\ = \; & \det\big((\lambda - A)(I - (\lambda - A)^{-1}tuu^*H)\big) \\ = \; & \det(\lambda - A)\det(I - (\lambda - A)^{-1}tuu^*H) \\ = \; & \det(\lambda - A)(1 - tu^*H(\lambda - A)^{-1}u) \\ = \; & (\lambda^2+b^2)^3(1 - tu^*H(\lambda - A)^{-1}u) \\ = \; & (\lambda^2+b^2)^3(1 - t\begin{bmatrix} -u_3^* & u_2^* - u_1^* \end{bmatrix}(\lambda - A)^{-1}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}) \\ = \; & (\lambda^2+b^2)^3(1 - t\left\{-u_3^*(\lambda-\gamma)^{-1}u_1 + u_2^*(\lambda-\gamma)^{-1}u_2 - u_1^*(\lambda-\gamma)^{-1}u_3\right\} \\ - \; & t\left\{-u_3^*(\lambda-\gamma)^{-2}u_2 + u_2^*(\lambda-\gamma)^{-2}u_3\right\} \\ + \; & tu_3^*(\lambda-\gamma)^{-3}u_3). \end{aligned}$$

Inserting the formulas for $(\lambda - \gamma)^{-j}$ for $j = 1, 2, 3$, we obtain from this that

$$\begin{aligned} & \det(\lambda I - (A + tuu^*H)) \\ = \; & (\lambda^2+b^2)^3 - t(\lambda^2+b^2)^2\left\{-u_3^*(\lambda+\gamma)u_1 + u_2^*(\lambda+\gamma)u_2 - u_1^*(\lambda+\gamma)u_3\right\} \\ - \; & t(\lambda^2+b^2)\left\{-u_3^*(\lambda^2+2\lambda\gamma+\gamma^2)u_2 + u_2^*(\lambda^2+2\lambda\gamma+\gamma^2)u_3\right\} \\ + \; & tu_3^*(\lambda^3+3\lambda^2\gamma+3\lambda\gamma^2+\gamma^3)u_3. \end{aligned}$$

Next, we use the fact that γ is skew-symmetric and that $\gamma^2 = -b^2 I_2$. This means that for any vectors $x, y \in \mathbb{R}^2$ we have $x^*\gamma x = 0$ and $x^*\gamma y + y^*\gamma x = 0$. Thus quite a number of terms in the above expression cancel, and we are left with

$$\begin{aligned} & \det(\lambda I - (A + tuu^*H)) \\ = \; & (\lambda^2+b^2)^3 - t(\lambda^2+b^2)^2\lambda\left\{-2u_3^*u_1 + u_2^*u_2\right\} \\ - \; & t(\lambda^2+b^2)\left\{-4\lambda u_3^*\gamma u_2\right\} + tu_3^*(\lambda^3 - 3b^2\lambda)u_3) \\ = \; & (\lambda^2+b^2)^3 - \\ & t\lambda\left\{(\lambda^2+b^2)^2(-2u_3^*u_1 + u_2^*u_2) + (\lambda^2+b^2)(-4u_3^*\gamma u_2 - u_3^*u_3) + 4b^2u_3^*u_3\right\}. \end{aligned} \tag{4.5}$$

Let us denote the polynomial in braces by $r(\lambda)$, so

$$r(\lambda) = (\lambda^2 + b^2)^2(-2u_3^*u_1 + u_2^*u_2) + (\lambda^2 + b^2)(-4u_3^*\gamma u_2 - u_3^*u_3) + 4b^2u_3^*u_3. \tag{4.6}$$

Now we are in a position to show that the theorem holds for this example. Indeed, if $\lambda = iy$ are pure-imaginary, then the terms $(\lambda^2 + b^2)^3$ are real, and so is $r(\lambda)$. If iy would be an eigenvalue of $A + tuu^*H$, then we would have

$$(b^2 - y^2)^3 - ityr(iy) = 0,$$

and since both $r(iy)$ and $b^2 - y^2$ are real, this would mean that they are both zero. Hence $y = \pm b$. This gives a contradiction.

Thus, for this example all eigenvalues of $A + tuu^*H$, which are not eigenvalues of A, do not lie on the imaginary line and this holds for all $t > 0$.

It remains to prove Parts 1 and 2 in Theorem 4.3.1. We start with a lemma which discusses the case where there is just one eigenvalue.

Lemma 4.3.6 *Suppose the H-positive real $n \times n$ matrix A has one eigenvalue (say $\hat{\lambda}$) and A is in Jordan form, i.e.*

$$A = J_{n_1}(\hat{\lambda}) \oplus \ldots \oplus J_{n_k}(\hat{\lambda})$$

with $k \geq 2$ and $n_1 \geq n_2 \geq \ldots \geq n_k$. Then for generic vectors u and all but finitely many $t > 0$, $\hat{\lambda}$ is an eigenvalue of

$$B(t) = A + tuu^*H.$$

In this case $B(t)$ has partial multiplicities $n_2 \geq \ldots \geq n_k$ with respect to the eigenvalue $\hat{\lambda}$ and has n_1 simple eigenvalues which are not eigenvalues of A.

Proof. To show that $\hat{\lambda}$ is an eigenvalue of $B(t) = A + tuu^*H$ for a generic set of vectors u and to determine the partial multiplicities of $B(t)$ at $\hat{\lambda}$ in this case, we simplify the calculations by replacing $B(t)$ by the matrix $B(t) - \hat{\lambda}I$ (and thus $\hat{\lambda}$ by 0).

Write H in the form

$$H = \begin{bmatrix} H_{11} & H_{12} & \ldots & H_{1k} \\ H_{21} & H_{22} & \ldots & H_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ H_{k1} & H_{k2} & \ldots & H_{kk} \end{bmatrix},$$

where H_{kl} is of dimension $n_k \times n_l$.

By (2.1) each diagonal block H_{ll} is an anti-triangular matrix of the form

$$H_{ll} = \begin{bmatrix} & & & -\epsilon_l \\ & 0 & \epsilon_l & \\ & \iddots & & \\ & \epsilon_l & * & \\ -\epsilon_l & & & \end{bmatrix}, \text{ for some } \epsilon_l \neq 0$$

and all off-diagonal blocks H_{lk} are anti-triangular of the form

$$H_{lk} = \begin{bmatrix} & & & \epsilon_{l_k} \\ & 0 & * & \\ & \iddots & & \\ & * & * & \\ \nu_{k_l} & & & \end{bmatrix}, \text{ for some } \epsilon_{l_k} \text{ and } \nu_{k_l}.$$

Here ϵ_{l_k} denotes the $(1, n_k)$-th entry of the block H_{lk} and ν_{k_l} denotes the $(n_l, 1)$-th entry of H_{lk}.

We have $n = \sum_{l=1}^{k} n_l$. Let

$$u = \begin{bmatrix} \underline{u}^{(1)} \\ \underline{u}^{(2)} \\ \vdots \\ \underline{u}^{(k)} \end{bmatrix}, \text{ with } \underline{u}^{(l)} = \begin{bmatrix} u^{(l,1)} \\ u^{(l,2)} \\ \vdots \\ u^{(l,n_l)} \end{bmatrix},$$

where $u^{(l,m)} \in \mathbb{C}$ for $l = 1, \ldots, k$ and $m = 1, \ldots, n_l$. Similarly, let

$$v = \begin{bmatrix} \underline{v}^{(1)} \\ \underline{v}^{(2)} \\ \vdots \\ \underline{v}^{(k)} \end{bmatrix}, \text{ with } \underline{v}^{(l)} = \begin{bmatrix} v^{(l,1)} \\ v^{(l,2)} \\ \vdots \\ v^{(l,n_l)} \end{bmatrix}.$$

Since $v^* = u^* H$ for H-positive real rank-one perturbations, we can write

$$v = Hu = \begin{bmatrix} \sum_{r=1}^{k} H_{1r} \underline{u}^{(r)} \\ \sum_{r=1}^{k} H_{2r} \underline{u}^{(r)} \\ \vdots \\ \sum_{r=1}^{k} H_{kr} \underline{u}^{(r)} \end{bmatrix}$$

i.e. $\underline{v}^{(l)} = \sum_{r=1}^{k} H_{lr} \underline{u}^{(r)}$ for $l = 1, \ldots, k$ and H_{lr} and H_{ll} as before. Thus, $\underline{v}^{(l)}$ is of the form

$$\underline{v}^{(l)} = \begin{bmatrix} \epsilon_{l_1} u^{(1,n_1)} \\ * \\ \vdots \\ * \end{bmatrix} + \begin{bmatrix} \epsilon_{l_2} u^{(2,n_2)} \\ * \\ \vdots \\ * \end{bmatrix} + \ldots + \begin{bmatrix} -\epsilon_l \, u^{(l,n_l)} \\ * \\ \vdots \\ * \end{bmatrix} + \ldots + \begin{bmatrix} \epsilon_{l_k} u^{(k,n_k)} \\ * \\ \vdots \\ * \end{bmatrix}.$$

This can be written as

$$\underline{v}^{(l)} = \begin{bmatrix} \sum_{\substack{r=1 \\ r \neq l}}^{k} \epsilon_{l_r} u^{(r,n_r)} - \epsilon_l \, u^{(l,n_l)} \\ * \\ \vdots \\ * \end{bmatrix}.$$

Let $S^{-1} = \mathrm{Toep}(\underline{v}^{(1)}) \oplus \mathrm{Toep}(\underline{v}^{(2)}) \oplus \ldots \oplus \mathrm{Toep}(\underline{v}^{(k)})$, where $\mathrm{Toep}(\underline{v}^{(l)})$ denotes the upper triangular Toeplitz matrix with entries on the first row given by $\underline{v}^{(l)T}$. We assume generically that $v^{(l,1)} \neq 0$ for $l = 1, \ldots, k$. This implies that S is indeed invertible. Now consider $S^{-1}B(t)S$, that is

$$\begin{aligned} S^{-1}B(t)S &= S^{-1}(A + tuu^*H)S \\ &= S^{-1}AS + tS^{-1}uu^*HS \\ &= A + tS^{-1}uu^*HS, \end{aligned}$$

see [43] Section 12.4. Recall that $v^* = u^*H$, hence $u^*HS = v^*S$. If we let $b^* = v^*S$, then

$$\begin{aligned} S^{-1}B(t)S &= A + tS^{-1}ub^* \\ &= \bigoplus_{r=1}^{k} J_{n_r}(\hat{\lambda}) + t \begin{bmatrix} \underline{\tilde{u}}^{(1)} \\ \underline{\tilde{u}}^{(2)} \\ \vdots \\ \underline{\tilde{u}}^{(k)} \end{bmatrix} b^*, \end{aligned}$$

where $S^{-1}u = [\underline{\tilde{u}}^{(1)}, \underline{\tilde{u}}^{(2)}, \ldots, \underline{\tilde{u}}^{(k)}]^*$. One verifies that $b^* = [1 \ldots 0 | 1 \ldots 0 | \ldots | 1 \ldots 0]$, hence the matrix $S^{-1}B(t)S$ has the form:

$$S^{-1}B(t)S = \left[\begin{array}{cccc|cccc|c|ccc} t\tilde{u}^{(1,1)} + \hat{\lambda} & 1 & \ldots & 0 & t\tilde{u}^{(1,1)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(1,1)} & 0 & \ldots & 0 \\ t\tilde{u}^{(1,2)} & \hat{\lambda} & \ldots & 0 & t\tilde{u}^{(1,2)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(1,2)} & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\ t\tilde{u}^{(1,n_1)} & 0 & \ldots & \hat{\lambda} & t\tilde{u}^{(1,n_1)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(1,n_1)} & 0 & \ldots & 0 \\ \hline t\tilde{u}^{(2,1)} & 0 & \ldots & 0 & t\tilde{u}^{(2,1)} + \hat{\lambda} & 1 & \ldots & 0 & \ldots & t\tilde{u}^{(2,1)} & 0 & \ldots & 0 \\ t\tilde{u}^{(2,2)} & 0 & \ldots & 0 & t\tilde{u}^{(2,2)} & \hat{\lambda} & \ldots & 0 & \ldots & t\tilde{u}^{(2,2)} & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\ t\tilde{u}^{(2,n_2)} & 0 & \ldots & 0 & t\tilde{u}^{(2,n_2)} & 0 & \ldots & \hat{\lambda} & \ldots & t\tilde{u}^{(2,n_2)} & 0 & \ldots & 0 \\ \hline & \downarrow & & & & \downarrow & & & \ddots & & \downarrow & & \\ \hline t\tilde{u}^{(k,1)} & 0 & \ldots & 0 & t\tilde{u}^{(k,1)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(k,1)} + \hat{\lambda} & 1 & \ldots & 0 \\ t\tilde{u}^{(k,2)} & 0 & \ldots & 0 & t\tilde{u}^{(k,2)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(k,2)} & \hat{\lambda} & \ldots & 0 \\ \vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\ t\tilde{u}^{(k,n_k)} & 0 & \ldots & 0 & t\tilde{u}^{(k,n_k)} & 0 & \ldots & 0 & \ldots & t\tilde{u}^{(k,n_k)} & 0 & \ldots & \hat{\lambda} \end{array}\right].$$

(Take note that each diagonal block of the matrix $S^{-1}B(t)S$ has ones on the superdiagonal.)

Now we consider the matrix:

$$\hat{B}(t) = S^{-1}B(t)S - \hat{\lambda}I =$$

$$\left[\begin{array}{cccc|cccc|c|cccc}
t\tilde{u}^{(1,1)} & 1 & \dots & 0 & t\tilde{u}^{(1,1)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(1,1)} & 0 & \dots & 0 \\
t\tilde{u}^{(1,2)} & 0 & \dots & 0 & t\tilde{u}^{(1,2)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(1,2)} & 0 & \dots & 0 \\
\vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\
t\tilde{u}^{(1,n_1)} & 0 & \dots & 0 & t\tilde{u}^{(1,n_1)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(1,n_1)} & 0 & \dots & 0 \\
\hline
t\tilde{u}^{(2,1)} & 0 & \dots & 0 & t\tilde{u}^{(2,1)} & 1 & \dots & 0 & \dots & t\tilde{u}^{(2,1)} & 0 & \dots & 0 \\
t\tilde{u}^{(2,2)} & 0 & \dots & 0 & t\tilde{u}^{(2,2)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(2,2)} & 0 & \dots & 0 \\
\vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\
t\tilde{u}^{(2,n_2)} & 0 & \dots & 0 & t\tilde{u}^{(2,n_2)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(2,n_2)} & 0 & \dots & 0 \\
\hline
 & \downarrow & & & & \downarrow & & & \ddots & & \downarrow & & \\
\hline
t\tilde{u}^{(k,1)} & 0 & \dots & 0 & t\tilde{u}^{(k,1)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(k,1)} & 1 & \dots & 0 \\
t\tilde{u}^{(k,2)} & 0 & \dots & 0 & t\tilde{u}^{(k,2)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(k,2)} & 0 & \dots & 0 \\
\vdots & & \ddots & & \vdots & & \ddots & & & \vdots & & \ddots & \\
t\tilde{u}^{(k,n_k)} & 0 & \dots & 0 & t\tilde{u}^{(k,n_k)} & 0 & \dots & 0 & \dots & t\tilde{u}^{(k,n_k)} & 0 & \dots & 0
\end{array}\right].$$

From the above matrix representation of $\hat{B}(t)$ it is easily verified that $\hat{B}(t)(e_1 - e_{n_1+1}) = 0$, i.e. 0 is an eigenvalue of $\hat{B}(t)$ and hence $\hat{\lambda}$ is an eigenvalue of $B(t)$ in the case when the vectors u are generically chosen such that $v^{(l,1)} \neq 0$ for $l = 1, \dots, k$. For the eigenvalue $\hat{\lambda}$, we find Jordan chains (for the matrix $S^{-1}B(t)S$) as

$$\begin{aligned}
&e_1 - e_{n_1+1}, e_2 - e_{n_1+2}, \dots, e_{n_2} - e_{n_1+n_2} \\
&e_1 - e_{n_1+n_2+1}, e_2 - e_{n_1+n_2+2}, \dots, e_{n_3} - e_{\sum_{r=1}^{3} n_r} \\
&\vdots \\
&e_1 - e_{(\sum_{r=1}^{k-1} n_r)+1}, e_2 - e_{(\sum_{r=1}^{k-1} n_r)+2}, \dots, e_{n_k} - e_{\sum_{r=1}^{k} n_r}.
\end{aligned} \tag{4.7}$$

There are thus at least $(k-1)$ chains of lengths $n_2 \geq n_3 \geq \dots \geq n_k$, respectively. Note that in the chains above, the vectors

$$e_1 - e_{(\sum_{r=1}^{l} n_r)+1} \quad \text{for } l = 1, 2, \dots, k-1$$

are eigenvectors of $S^{-1}B(t)S$ associated with the eigenvalue $\hat{\lambda}$. This shows that

$$\dim(\mathrm{Ker}(S^{-1}B(t)S - \hat{\lambda}I)) \geq k - 1.$$

Recall that if $\hat{B}(t) = S^{-1}B(t)S - \hat{\lambda}I$, then the characteristic polynomial of $\hat{B}(t)$ is given by

$$\begin{aligned}
p_{\hat{B}(t)}(\lambda) &= \det(\hat{B}(t) - \lambda I) \\
&= \det((S^{-1}B(t)S - \hat{\lambda}I) - \lambda I) \\
&= \det(B(t) - \hat{\lambda}I - \lambda I) \\
&= \det(A + tuu^*H - \hat{\lambda}I - \lambda I) \\
&= \det((A - \hat{\lambda}I) - \lambda I + tuu^*H) \\
&= \det([(A - \hat{\lambda}I) - \lambda I][I + t((A - \hat{\lambda}I) - \lambda I)^{-1}uu^*H]), \text{ for } \lambda \neq 0 \\
&= \det((A - \hat{\lambda}I) - \lambda I)(1 + tu^*H((A - \hat{\lambda}I) - \lambda I)^{-1}u) \qquad \text{(Lemma 4.6.1)} \\
&= \det((A - \hat{\lambda}I) - \lambda I) + tu^*H[\det(A - (\hat{\lambda} + \lambda)I)]((A - \hat{\lambda}I) - \lambda I)^{-1}u \\
&= p_{A-\hat{\lambda}I}(\lambda) + tq(\lambda).
\end{aligned}$$

Since we have on both sides of the equation polynomials, it also holds for $\lambda = 0$. Here $p_{A-\hat{\lambda}I}(\lambda)$ is the characteristic polynomial of degree n. Below we verify that $q(\lambda)$ is a polynomial of degree $\leq n - 1$.

If $||A - \hat{\lambda}I|| < |\lambda|$, then we may rewrite $[(A - \hat{\lambda}I) - \lambda I]^{-1}$ as

$$\begin{aligned}
[(A - \hat{\lambda}I) - \lambda I]^{-1} &= [\lambda(\lambda^{-1}(A - \hat{\lambda}I) - I)]^{-1} \\
&= -\lambda^{-1}(I - \lambda^{-1}(A - \hat{\lambda}))^{-1} \\
&= -\lambda^{-1}\sum_{j=0}^{\infty}(\lambda^{-1}(A - \hat{\lambda}I))^j \\
&= -\sum_{j=0}^{\infty}\frac{(A - \hat{\lambda}I)^j}{\lambda^{j+1}}.
\end{aligned}$$

Now,

$$A - \hat{\lambda}I = J_{n_1}(0) \oplus J_{n_2}(0) \oplus \ldots \oplus J_{n_k}(0).$$

Thus, $(A - \hat{\lambda}I)^{n_1} = 0$ and $(A - \hat{\lambda}I)^{n_1-1}u \neq 0$ for generic $u \neq 0$. If $n_1 > n_2$, then

$$(A - \hat{\lambda}I)^{n_1-1}u = \left[\begin{array}{cccc|c} 0 & 0 & \ldots & 1 & \\ 0 & 0 & \ldots & 0 & \\ & & \ddots & & \\ 0 & 0 & \ldots & 0 & \\ \hline & & & & \ddots \end{array}\right] \begin{bmatrix} \underline{u}^{(1)} \\ \vdots \\ \underline{u}^{(k)} \end{bmatrix} = \left[\begin{array}{c} u^{(1,n_1)} \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \\ \hline \vdots \end{array}\right].$$

(If $n_1 = n_2$, then $(A - \hat{\lambda}I)^{n_1-1}u = [u^{(1,n_1)}\,0\ldots0|u^{(2,n_2)}\,0\ldots0|\ldots]^*$, etc.) Hence,

$$\begin{aligned} q(\lambda) &= -u^*H(\sum_{j=0}^{n_1-1}\frac{(A-\hat{\lambda}I)^j}{\lambda^{j+1}})u\det(A-\hat{\lambda}I-\lambda I) \\ &= -((-1)^n\lambda^n)u^*H(\sum_{j=0}^{n_1-1}\frac{(A-\hat{\lambda}I)^j}{\lambda^{j+1}})u \\ &= (-1)^{n+1}u^*H(\sum_{j=0}^{n_1-1}\lambda^{n-j-1}(A-\hat{\lambda}I)^j)u. \end{aligned}$$

Thus we have

$$\begin{aligned} p_{\hat{B}(t)}(\lambda) &= p_{A-\hat{\lambda}I}(\lambda)+(-1)^{n+1}(tu^*H\sum_{j=0}^{n_1-1}\lambda^{n-j-1}(A-\hat{\lambda}I)^j)u \\ &= (-1)^n\lambda^n+(-1)^{n+1}tu^*H(\sum_{j=0}^{n_1-1}\lambda^{n-j-1}(A-\hat{\lambda}I)^j)u \\ &= (-1)^n\lambda^{n-n_1}[\lambda^{n_1}-tu^*H(\sum_{j=0}^{n_1-1}\lambda^{n_1-j-1}(A-\hat{\lambda}I)^j)u]. \end{aligned} \tag{4.8}$$

We may rewrite (4.8) as:

$$p_{\hat{B}(t)}(\lambda)=\lambda^{n-n_1}((-1)^n\lambda^{n_1}+\alpha_{n_1-1}\lambda^{n_1-1}+\ldots+\alpha_1\lambda+\alpha_0), \tag{4.9}$$

where we have

$$\begin{aligned} \alpha_0 &= (-1)^{n+1}tu^*H(A-\hat{\lambda}I)^{n_1-1}u \\ \alpha_1 &= (-1)^{n+1}tu^*H(A-\hat{\lambda}I)^{n_1-2}u \\ \alpha_2 &= (-1)^{n+1}tu^*H(A-\hat{\lambda}I)^{n_1-3}u \\ &\vdots \end{aligned}$$

In general

$$\alpha_j=(-1)^{n+1}tu^*H(A-\hat{\lambda}I)^{n_1-j-1}u,\quad j=0,\ldots,n_1-1.$$

For $\alpha_0\neq 0$, we must have

$$(A-\hat{\lambda}I)^{n_1-1}u\neq 0.$$

This is the case when the coordinate $u^{(1,n_1)}$ of the vector $\underline{u}^{(1)}$ is non-zero, i.e. the choice of all $u\in\mathbb{R}^n$ such that $(A-\hat{\lambda}I)^{n_1-1}u\neq 0$ gives a generic set.

It is then clear that $\hat{B}(t)$ has an eigenvalue of algebraic multiplicity $n-n_1$ at zero, or equivalently, that $B(t)$ has an eigenvalue of algebraic multiplicity $n-n_1=n_2+n_3+\ldots+n_k$ at $\hat{\lambda}$.

To show that the eigenvalues of $\hat{B}(t)$, different from those of A are simple, we consider (4.9) again. It suffices to show that the polynomial $h(\lambda) = (-1)^n\lambda^{n_1} + \alpha_{n_1-1}\lambda^{n_1-1} + \ldots + \alpha_1\lambda + \alpha_0$ has n_1 different roots. To show this we slightly adapt an argument from [68]. Write $\alpha_j = t\beta_j$ with

$$\beta_j = (-1)^{n+1}u^*H(A - \hat{\lambda}I)^{n_1-j-1}u.$$

Then $h(\lambda) = 0$ if and only if

$$(-1)^n\lambda^{n_1} + t\beta_{n_1-1}\lambda^{n_1-1} + \ldots + t\beta_1\lambda + t\beta_0 = 0.$$

Dividing through by t, this is zero if and only if

$$k(\lambda, t) := (-1)^n\lambda^{n_1}/t + \beta_{n_1-1}\lambda^{n_1-1} + \ldots + \beta_1\lambda + \beta_0 = 0.$$

Introduce $y = t^{1/n_1}$ and $x = \frac{\lambda}{y}$, then

$$k(\lambda, t) = P(x, y) = (-1)^n x^{n_1} + \beta_{n_1-1}(xy)^{n_1-1} + \ldots + \beta_1 xy + \beta_0.$$

We now see that $P(x, 0) = (-1)^n x^{n_1} + \beta_0 = 0$ with $\beta_0 \neq 0$ has n_1 distinct nonzero roots. Now use the implicit function theorem, see e.g. [15], Section 10.2, in particular (10.2.4) to deduce that there are n_1 distinct functions $x_j(y)$, $j = 1, \ldots, n_1$, analytic in y for $y \neq 0$ small enough, such that $P(x_j(y), y) = 0$. Thus $\hat{B}(t)$ has n_1 distinct eigenvalues for $t > 0$ small enough. To complete the proof for all but finite many $t > 0$, we use the argument from [65] Theorem 6.9. See also the final part of the proof of Theorem 4.3.1 below. □

Remark 4.3.7 (1) Note that the proof of Lemma 4.3.6 actually proves Part 1 and 2 of Theorem 4.3.1 for the case of a single eigenvalue.
(2) The eigenvalues of $B(t)$ which are not eigenvalues of A, close to $\lambda_0 \in \sigma(A)$ behave as the roots of $(-1)^n\lambda^{n_1} + \alpha_0$. (Recall that $\alpha_0 = t\beta_0$.)

Proof of Part (1) and (2) of Theorem 4.3.1. Without loss of generality, let A be in Jordan normal form as before:

$$A = \bigoplus_{j=1}^{p}\bigoplus_{i=1}^{k_j} J_{n_j,i}(\lambda_j).$$

We consider a rank-one perturbation $B(t) = A + tuu^*H$ of A as before. The proof of Lemma 4.3.6 can be used to prove that $B(t)$ generically has k_{j-1} Jordan chains of lengths $n_{j,2} \geq n_{j,3} \geq \ldots \geq n_{j,k_j}$ corresponding to the eigenvalue λ_j.

The proof of the general case now proceeds as the proof of Lemma 4.3.6 by considering the characteristic polynomial of $B(t)$.

$$\begin{aligned} p_{B(t)}(\lambda) &= p_A(\lambda)(1 - tu^*H(\lambda I - A)^{-1}u) \\ &= \prod_{j=1}^{p}\prod_{i=2}^{k_j}(\lambda - \lambda_j)^{n_{j,i}}[(m_A(\lambda) - tu^*m_A(\lambda)H(\lambda I - A)^{-1}u)], \end{aligned}$$

where $m_A(\lambda) = \prod_{j=1}^{p}(\lambda - \lambda_j)^{n_{j,1}}$ is the *minimal polynomial*. Since we are interested in those eigenvalues of $B(t)$ which are not eigenvalues of A (i.e., the set $\sigma(B(t))\backslash\sigma(A)$), and since these eigenvalues are the roots of

$$q(\lambda) := m_A(\lambda) - tu^* m_A(\lambda) H(\lambda I - A)^{-1} u,$$

we only consider the latter polynomial. Let us, for sake of simplicity of notation, replace the subscript $n_{j,1}$ by n_j. Then $q(\lambda)$ can be expressed as

$$q(\lambda) = \prod_{j=1}^{p}(\lambda - \lambda_j)^{n_j}(1 - tu^* H(\lambda I - A)^{-1} u).$$

Since $\lambda \neq \lambda_j$, we write

$$(\lambda I - A)^{-1} = \sum_{j=1}^{p}\sum_{m=1}^{n_j}(\lambda - \lambda_j)^{-m}(A - \lambda_j I)^{m-1} P_j,$$

where P_j is the spectral projection onto the eigenspace of the eigenvalue λ_j, cf. Chapter 2.1 of [23].

Since the polynomial $q(\lambda)$ above also depends on $t > 0$, we agree to denote it by $q(\lambda, t)$ for the moment. Substituting the above expression for $(\lambda I - A)^{-1}$ into $q(\lambda)$ we then have

$$\begin{aligned} q(\lambda) &= m_A(\lambda) - t\sum_{j=1}^{p}\Big(\prod_{i\neq j}^{p}(\lambda - \lambda_i)^{n_i}\Big)\Big(\sum_{m=1}^{n_j}(\lambda - \lambda_j)^{n_j - m} u^* H P_j (A - \lambda_j I)^{m-1} P_j u\Big) \quad (4.10) \\ &= (\lambda - \lambda_1)^{n_1} - t\prod_{i=2}^{p}(\lambda - \lambda_i)^{n_i}\Big(\sum_{m=1}^{n_1}(\lambda - \lambda_1)^{n_1 - m} u^* H P_1 (A - \lambda_1 I)^{m-1} P_1 u\Big) \\ &- t\sum_{j=2}^{p}\Big(\prod_{i\neq j}^{p}(\lambda - \lambda_j)^{n_i}\Big(\sum_{m=1}^{n_j}(\lambda - \lambda_j)^{n_j - m} u^* H P_j (A - \lambda_j I)^{m-1} P_j u\Big) \\ &= \prod_{j=2}^{p}(\lambda - \lambda_j)^{n_j} q_1(\lambda, t) - t\hat{r}(\lambda)(\lambda - \lambda_1)^{n_1}, \quad (4.11) \end{aligned}$$

where we put

$$q_1(\lambda, t) = (\lambda - \lambda_1)^{n_1} - t\sum_{m=1}^{n_1}(\lambda - \lambda_1)^{n_1 - m} u^* H P_1 (A - \lambda_1 I)^{m-1} P_1 u$$

and

$$t\hat{r}(\lambda)(\lambda - \lambda_1)^{n_1} = q(\lambda, t) - \prod_{j=2}^{p}(\lambda - \lambda_j)^{n_j} q_1(\lambda, t).$$

Introduce $x = t^{1/n_1}$ and $y = \frac{\lambda - \lambda_1}{x}$, then

$$\begin{aligned} q(\lambda,t) &= \prod_{j=2}^{p}(xy+\lambda_1-\lambda_j)^{n_j}[(xy)^{n_1}-x^{n_1}\sum_{m=1}^{n_1}(xy)^{n_1-m}u^*HP_1(A-\lambda_1)^{m-1}P_1u] \\ &- x^{2n_1}y^{n_1}\hat{r}(x,y). \end{aligned}$$

Observe that the right hand side is a polynomial in x and y, and that all terms contain x^{n_1} as a factor. Divide by x^{n_1} and introduce

$$\begin{aligned} P(x,y) &= \prod_{j=2}^{p}(xy+\lambda_1-\lambda_j)^{n_j}[y^{n_1}-\sum_{m=1}^{n_1-1}(xy)^{n_1-m}u^*HP_1(A-\lambda_1)^{m-1}P_1u \\ &- u^*HP_1(A-\lambda_1)^{n_1-1}P_1u]-x^{n_1}y^{n_1}\hat{r}(x,y). \end{aligned}$$

Then

$$P(0,y)=\prod_{j=2}^{p}(\lambda_1-\lambda_j)^{n_j}[y^{n_1}-u^*HP_1(A-\lambda_1)^{n_1-1}P_1u].$$

Denoting $P_1u=u_1=\begin{bmatrix}u_{11} & \dots & u_{1n_1} & 0 & \dots & 0\end{bmatrix}^*$ and taking (A,H) in simple form, as in Lemma 4.3.6, we have

$$(A-\lambda_1)^{n_1-1}P_1u=\begin{bmatrix}u_{1n_1}\\0\\\vdots\\0\end{bmatrix}=u_{1n_1}e_1$$

and

$$u^*HP_1=\begin{bmatrix}u_{11} & \dots & u_{1n_1} & 0 & \dots & 0\end{bmatrix}\left[\begin{array}{ccc|c} 0 & \dots & \overline{a} & \\ \vdots & \iddots & * & 0 \\ a & \dots & & \\ \hline & 0 & & 0 \end{array}\right].$$

So, $u^*HP_1(A-\lambda_1I)^{n_1-1}P_1u=au_{1n_1}^2$ if $\lambda_1\in\mathbb{R}$. Thus

$$P(0,y)=\prod_{j=2}^{p}(\lambda_1-\lambda_j)^{n_j}[y^{n_1}-au_{1n_1}^2].$$

Now by applying the analytic version of the implicit function theorem, see p 272 of [15], we see that $P(x,y)=0$ has n_1 solutions, namely, $y_j(x)$ for $j=1,\dots,n_1$, with

$$y_j(0)=a^{1/n_1}u_{1n_1}^{2/n_1}e^{\frac{2\pi i}{n_1}j}.$$

Translating to $\lambda_j(t)$, we have

$$\lambda_j(t)=\lambda_1+a^{1/n_1}u_{1n_1}^{2/n_1}e^{\frac{2\pi i}{n_1}j}t^{1/n_1}+\dots.$$

We are now in a position to finish the proof of Part 1 of Theorem 4.3.1 in the same way as was done in Lemma 4.3.6. Just counting multiplicities we see that the algebraic multiplicity of λ_1 as an eigenvalue of $B(t)$ is $n_{1,2} + \ldots + n_{1,k_1}$. Since the same arguments will apply to all eigenvalues, Part 1 is proved.

Also, Part 2 is proved for $t > 0$ small enough: It remains to prove that for all but finitely many $t > 0$ the eigenvalues of $B(t)$ which are not eigenvalues of A are simple eigenvalues. For this we use the argument from [65], Theorem 6.1. Recall that $m_A(\lambda) = \prod_{j=1}^{p}(\lambda - \lambda_j)^{n_{j,1}}$. Now let $p(\lambda) = u^* m_A(\lambda) H(\lambda I - A)^{-1} u$. If λ is a double eigenvalue of $B(t)$ and $\lambda \notin \sigma(A)$, then λ is a common root of $m_A(\lambda) - tp(\lambda)$ and $m_A'(\lambda) - tp'(\lambda)$.

Hence the *Sylvester resultant matrix* of these two polynomials must be singular (see e.g. [43] Theorem 13.3.2). As in [65] it may be seen that the determinant of the Sylvester resultant matrix is a polynomial in the variable t which is not identically zero. Hence, there is only a finite number of values $t > 0$ for which $B(t)$ has an eigenvalue of multiplicity at least two. □

An alternative approach to the one discussed above regarding the algebraic multiplicity of the eigenvalues of $p_{\hat{B}(t)}(\lambda)$, is to consider Theorem 2.3 in [55].

4.4 Asymptotics of eigenvalues

In this section we study the asymptotics of the eigenvalues of $A + tuu^*H$ as $t \to \infty$. We start by continuing the Example 4.3.5.

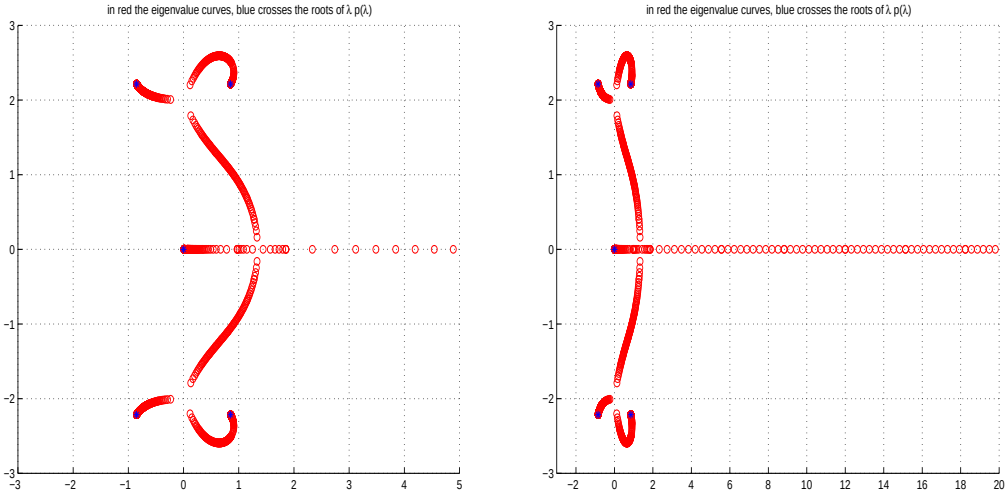
Example 4.4.1

In the situation of Example 4.3.5 we had computed that the eigenvalues of $A + tuu^*H$ were given by the roots of the polynomial

$$(\lambda^2 + b^2)^3 - t\lambda r(\lambda),$$

where $r(\lambda)$ is given by (4.6).

The next picture shows the eigenvalue curves for the specific vector u given by $u = \begin{bmatrix} -0.4326 & -1.6656 & 0.1253 & 0.2877 & -1.1465 & 1.1909 \end{bmatrix}^*$, which is obtained in matlab with `randn(6,1)`. As can be seen the eigenvalues converge to five points in the finite complex plane, given by 0, and $\pm(0.8507 \pm 2.2142i)$, and, although this can not be seen from the first picture, it can be seen from the second, one eigenvalue converges to plus infinity.



Applying Lemma 4.6.2 to $(\lambda^2+b^2)^3-t\lambda r(\lambda)$, we see that the location of the eigenvalues of $A+uu^*H$ as $t\to\infty$ will be at the roots of $\lambda r(\lambda)$, which are indeed the points 0, and $\pm(0.8507\pm 2.2142i)$.

We shall show that this is generically the typical behavior of the eigenvalues. Recall that the eigenvalues of $B(t)$ which are not eigenvalues of A are the roots of the polynomial

$$q(\lambda) = m_A(\lambda) - tu^*Hm_A(\lambda)(\lambda I - A)^{-1}u.$$

We denote

$$p(\lambda) = u^*Hm_A(\lambda)(\lambda I - A)^{-1}u,$$

and let l be the degree of q. Then the following proposition holds.

Proposition 4.4.2 *Let A be H-positive real, then for generic vectors $u\in\mathbb{C}^n$ the eigenvalues of $B(t)=A+tuu^*H$ will behave as follows: exactly $l-1$ of them approximate the roots of p, one goes to infinity, asymptotically along the half line tu^*Hu, $t>0$.*

Proof. The proof is immediate from Lemma 4.6.2. □

In the next example we consider the case where there are two blocks of the same size with zero eigenvalue. This example shows that it is possible that *asymptotically* the eigenvalues of $A+tuu^*H$ approach the imaginary axis.

Example 4.4.3

In this example we consider the case where

$$A=\begin{bmatrix} J_3(0) & 0\\ 0 & J_3(0)\end{bmatrix},\qquad H=\begin{bmatrix} H_1 & 0\\ 0 & H_1\end{bmatrix},$$

where $J_3(0)$ is the 3×3 Jordan block with zero eigenvalue, and

$$H_1=\begin{bmatrix} 0 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{bmatrix}.$$

In this case too, we compute $\det(\lambda I - (A + tuu^*H))$. With similar arguments as in Example 4.3.5, we obtain

$$\det(\lambda I - (A + tuu^*H)) = \lambda^3 \left(\lambda^3 - t\left\{\lambda^2(u_2^2 + u_5^2 - 2u_1u_3 - 2u_4u_6) - (u_3^2 + u_6^2)\right\}\right).$$

We see that as predicted by Theorem 4.3.1 there remain three eigenvalues at zero. Applying Lemma 4.6.2 with $p_1(\lambda) = \lambda^3$ and

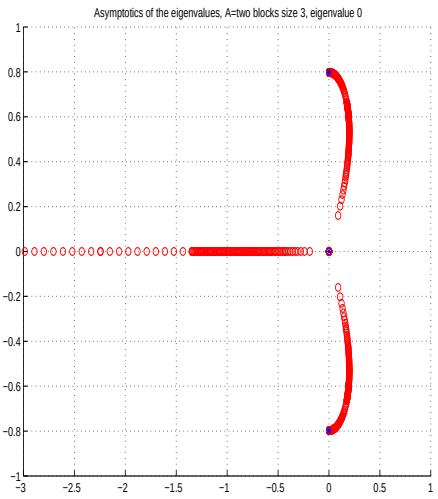
$$p_2(\lambda) = \lambda^2(u_2^2 + u_5^2 - 2u_1u_3 - 2u_4u_6) - (u_3^2 + u_6^2) = \lambda^2 u^*Hu - (u_3^2 + u_6^2)$$

we observe that the eigenvalues of $A + tuu^*H$ will be as follows: three at zero, one going to plus or minus infinity, and two converging to the roots of p_2. If u is H-positive, that is, $u^*Hu > 0$, one eigenvalue will go to $+\infty$, and the other two will approach two real numbers $\pm\sqrt{(u_3^2 + u_6^2)/u^*Hu}$. If u is H-negative, i.e. $u^*Hu < 0$, one eigenvalue will go to $-\infty$, and the other two will approach two pure imaginary numbers $\pm i\sqrt{-(u_3^2 + u_6^2)/u^*Hu}$. Also observe that $p_1 - tp_2$ again cannot have a pure-imaginary root, thus confirming Theorem 4.3.1.

The following figures show the eigenvalues of $B(t)$ as a function of $t > 0$.
Taking

$$u = \begin{bmatrix} 1.1892 & -0.0376 & 0.3273 & 0.1746 & -0.1867 & 0.7258 \end{bmatrix}^*$$

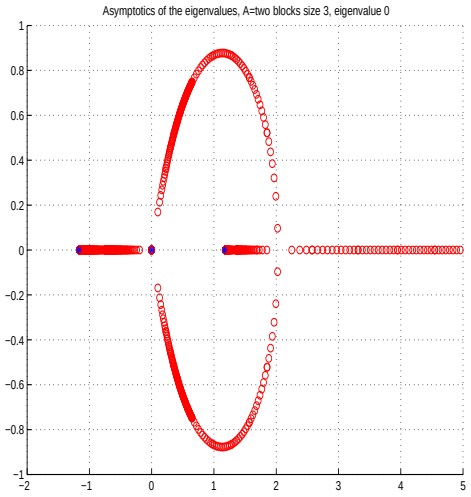
the roots of p are $\pm 0.7979i$.



Letting

$$u = \begin{bmatrix} -0.1858 & 0.0089 & 0.8369 & -0.7223 & -0.7215 & -0.2012 \end{bmatrix}^*$$

the roots of p are ± 1.1702.



As we have seen above, the qualitative behavior of the eigenvalues of $B(t)$ as a function of $t > 0$ is always as in either one of these two figures.

Example 4.4.4

The next example analyzes the case where we have

$$A = J_4(0) \oplus J_4(0), \qquad H = \begin{bmatrix} H_1 & 0 \\ 0 & H_1 \end{bmatrix}, \quad \text{and} \quad H_1 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \tag{4.12}$$

for different choices of u.

Again, as predicted by Theorem 4.3.1 there remain four eigenvalues at zero, one eigenvalue goes to plus or minus infinity, the other three approximate finite values. In this case, the characteristic polynomial of $A + tuu^*H$ is

$$\lambda^4 \left(\lambda^4 - t\left\{2\lambda^3(-u_1u_4 + u_2u_3 - u_5u_8 + u_6u_7) + \lambda^2(u_3^2 + u_7^2) - (u_4^2 + u_8^2)\right\}\right)$$

As before, applying Lemma 4.6.2 to $p_1(\lambda) = \lambda^4$ and

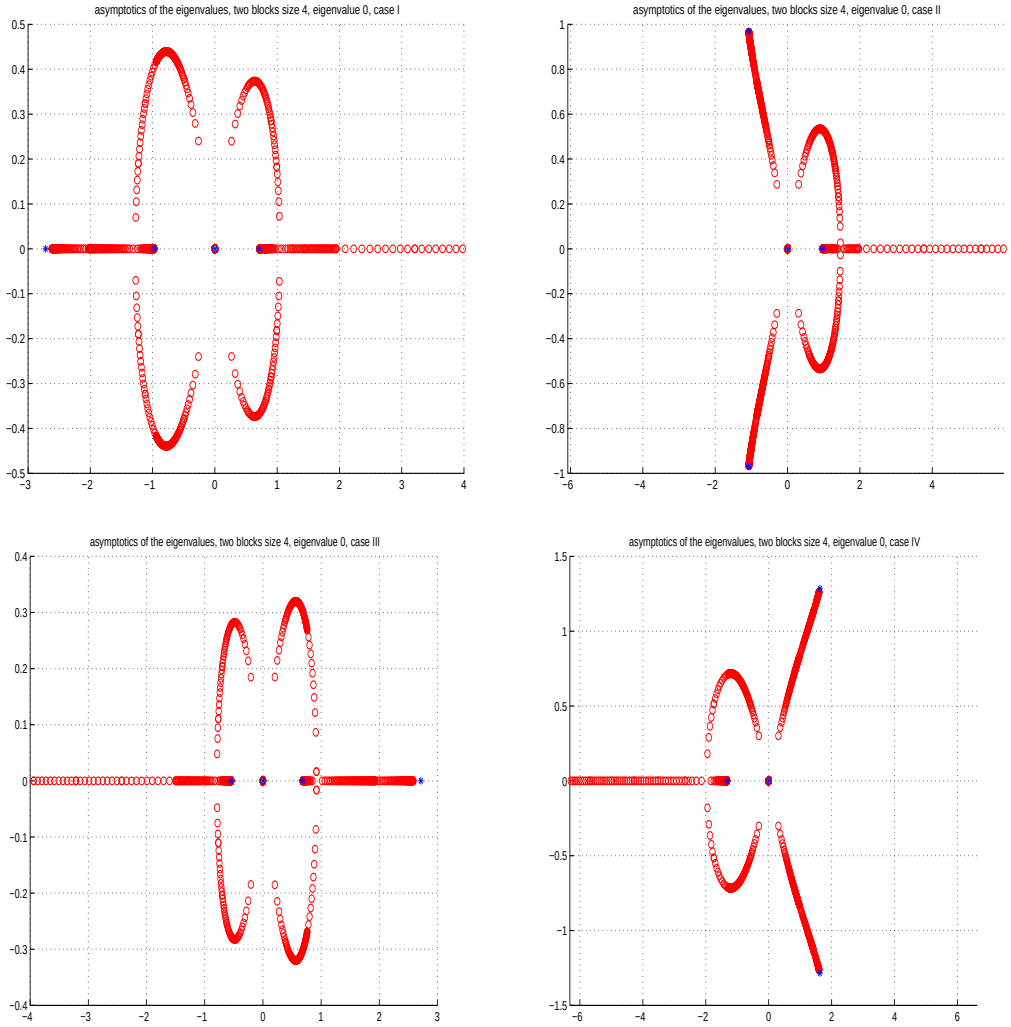
$$p_2(\lambda) = 2\lambda^3(-u_1u_4 + u_2u_3 - u_5u_8 + u_6u_7) + \lambda^2(u_3^2 + u_7^2) - (u_4^2 + u_8^2),$$

we see that the eigenvalues of $A + tuu^*H$ will be as follows: four at zero (one Jordan block of size four), one going to infinity, and three approaching the roots of p_2. Note that the leading coefficient of $p_2(\lambda)$ is u^*Hu. Hence, if $u^*Hu > 0$, two eigenvalues will go into the right half-plane and two into the left half-plane. The one eigenvalue to the right goes to positive infinity and the other one to the positive real root of $p_2(\lambda)$. The other two eigenvalues will approach either two negative real roots (see figure case I) or approach two conjugate complex roots in the left half-plane (see figure case II).

On the other hand, if $u^*Hu < 0$, again two eigenvalues go to the right half-plane and two eigenvalues to the left half-plane. The one eigenvalue to the left goes to minus infinity and the other one to a negative real root of $p_2(\lambda)$. The other two eigenvalues, to the right,

will approach either two real roots (see figure case III) or two conjugate complex roots in the right half-plane (see figure case IV).

Also note that this example as well confirms the result of Theorem 4.3.1



The four different cases when (A, H) is given by (4.12)

Example 4.4.5

The final example concerns a case where the eigenvalues are not on the imaginary line. Consider

$$A = J_4(-1 \pm 2i) \oplus J_3(3 \pm 4i), \qquad H = -I_8 \oplus I_6.$$

This is a pair of 14×14 matrices. We again take a 14×1 random vector u. As in the previous examples, let us denote $p_1(\lambda) = \det(\lambda - A)$, and

$$\det(\lambda - (A + tuu^*H)) = p_1(\lambda) - tp_2(\lambda).$$

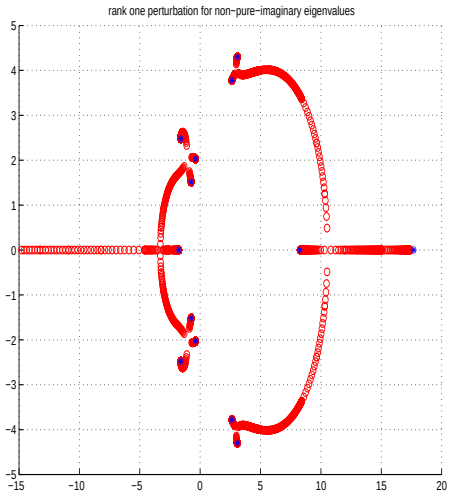
Then (with a little help from the symbolic toolbox of Matlab), one computes

$$\begin{aligned} p_2(\lambda) &= \frac{6288311797604141}{4503599627370496}\lambda^{13} - \frac{740779731774015}{17592186044416}\lambda^{12} + \\ &+ \frac{3538118983700225}{8796093022208}\lambda^{11} - \frac{2217132684997471}{1099511627776}\lambda^{10} + \\ &+ \frac{3566409675069751}{549755813888}\lambda^{9} - \frac{1411268841389411}{137438953472}\lambda^{8} + \\ &+ \frac{2463976178260191}{34359738368}\lambda^{7} - \frac{8027690789718847}{137438953472}\lambda^{6} + \\ &+ \frac{5675411788333239}{8589934592}\lambda^{5} + \frac{5306559348455229}{2147483648}\lambda^{4} + \\ &+ \frac{8911228193448319}{1073741824}\lambda^{3} + \frac{2527175811903299}{134217728}\lambda^{2} + \\ &+ \frac{3027100170486679}{134217728}\lambda + \frac{6126686202701393}{268435456}. \end{aligned}$$

(Obviously, because of the use of the symbolic toolbox, the fractions shown here are the best rational approximations to real numbers. The roots of p_2 can be computed by converting the symbolic expression to a numerical one, and then asking Matlab to compute the roots.) The roots of p_2 are given by the following three real numbers and five pairs of complex conjugate numbers

$$\begin{aligned} &17.6987, 8.2494, -1.7697, 3.0839 \pm 4.2992i, 2.6395 \pm 3.7781i, \\ &-1.6054 \pm 2.4782i, -0.3779 \pm 2.0289i, -0.7506 \pm 1.5171i. \end{aligned}$$

The following figure shows that again, the roots of p_2 are limits of eigenvalues of $A+tuu^*H$, when $t \to \infty$.



A question that may rise is: To what extent is all we did here on asymptotics going through in the general case, i.e., for $B(t) = A+tuv^*$, where A has no particular structure? This has been done in the work of [65].

4.5 Rank-one perturbations of H-positive real matrices - The real case

We conclude the main body of the chapter by making a few remarks regarding the real case. It is especially different at eigenvalue zero, if compared to the complex case. We will show this with two examples by considering the canonical form for the case when $HA+A^TH = 0$ and even dimension blocks at eigenvalue zero. (This can also be compared with what happened with rank-one perturbations for the real J-Hamiltonian matrices, see [55].) Before turning to the examples we remark that Theorem 4.3.1 is true for the real case as well when A is invertible.

Example 4.5.1

Let $A = J_2(0) \oplus J_2(0)$, and let H be given by

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Then $HA + A^TH = 0$. Let us consider a generic structured real H-nonnegative perturbation of rank-one of A: $B = A + uu^TH$. We denote $u = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix}^T$. First observe that a vector $x_0 = \begin{bmatrix} x_{01} & x_{02} & x_{03} & x_{04} \end{bmatrix}^T$ is in $\ker B$ if and only if $Ax = -uu^THx$. Now Ax has zero as its second and fourth coordinates, and as u is generic, we may assume that $u_2 \neq 0$ and $u_4 \neq 0$. That implies that both $Ax = 0$ and $u^THx = 0$. In particular, $x_{02} = x_{04} = 0$, and $-u_2x_{03} + u_4x_{01} = 0$. So without loss of generality we can take

$$x_0 = \begin{bmatrix} u_2 & 0 & u_4 & 0 \end{bmatrix}^T.$$

Next, we are looking for a vector x_1 such that $Bx_1 = x_0$. Denote $x_1 = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \end{bmatrix}^T$. Then $Bx_1 = Ax_1 + uu^THx_1 = x_0$. In the same way as in the previous paragraph we can now conclude that $u^THx_1 = 0$, and that $Ax_1 = x_0$. These two conditions combined imply, that we may put

$$x_1 = \begin{bmatrix} -u_1 & u_2 & -u_3 & u_4 \end{bmatrix}^T,$$

and adding any multiple of x_0 would work as well.

Further, we check whether it is possible that $Bx_2 = x_1$ for some vector x_2. Again, we denote $x_2 = \begin{bmatrix} x_{21} & x_{22} & x_{23} & x_{24} \end{bmatrix}^T$. Then

$$Bx_2 = Ax_2 + uu^THx_2 = \begin{bmatrix} x_{22} \\ 0 \\ x_{24} \\ 0 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} u^THx_2 = \begin{bmatrix} -u_1 \\ u_2 \\ -u_3 \\ u_4 \end{bmatrix}.$$

Comparing the second and fourth coordinate, we see that it is necessary that $u^THx_2 = 1$, and considering the first and third coordinate we note that we should take $x_{22} = -2u_1$, $x_{24} = -2u_3$. We arrive at the condition that

$$u^THx_2 = -x_{23}u_2 + x_{21}u_4 = 1.$$

Obviously, it is possible to arrange for this to hold by an appropriate choice of x_{23} or x_{21}, provided not both u_2 and u_4 are zero (which is generically the case).

We conclude that the rank-one perturbation B of A has a Jordan chain of length at least three corresponding to the zero eigenvalue. Since the trace of B is given by $\mathrm{tr}B = \mathrm{tr}A + \mathrm{tr}(uu^TH) = 0 + \mathrm{tr}(u^THu) = u^THu$, which is non-zero and positive for generic u, we have that B generically has one positive eigenvalue. Thus, the Jordan structure for a generic structured rank-one real perturbation is completely determined: $J_3(0) \oplus J_1(\lambda_0)$ with $\lambda_0 > 0$. Observe that this is quite different from the complex case, and that it is also quite different from the case of general unstructured rank-one perturbations.

Example 4.5.2

Let $A = J_2(0) \oplus J_2(0) \oplus J_2(0)$. Let us denote by Σ the matrix H from the previous example, and let P be the two by two sip matrix, that is $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Let $H = \Sigma \oplus P$. Then $HA + A^TH \geq 0$ (but not equal to zero). Again, let us consider a generic structured rank-one perturbation $B = A + uu^TH$.

First we observe that the dimension of the kernel of B is two. Then we check that the following vectors form two Jordan chains of length two:

$$x_{01} = \begin{bmatrix} -u_6 \\ 0 \\ 0 \\ 0 \\ u_4 \\ 0 \end{bmatrix}, \qquad x_{11} = \begin{bmatrix} -u_5 \\ -u_6 \\ 0 \\ 0 \\ -u_3 \\ u_4 \end{bmatrix}, \qquad x_{02} = \begin{bmatrix} 0 \\ 0 \\ u_6 \\ 0 \\ u_2 \\ 0 \end{bmatrix}, \qquad x_{12} = \begin{bmatrix} 0 \\ 0 \\ u_5 \\ u_6 \\ -u_1 \\ u_2 \end{bmatrix}.$$

So $Bx_{0i} = 0, Bx_{1i} = x_{0i}$, for $i = 1, 2$. Furthermore, one can check that there is no vector such that $By = x_{1i}$ for either $i = 1$ or $i = 2$. So these are two Jordan chains of length two. Moreover,

$$x_{01}^T H x_{11} = u_4^2 > 0, \qquad x_{02}^T H x_{12} = u_2^2 > 0.$$

So these blocks are uncoupled in the usual manner. Consider

$$\begin{aligned}
\det(A + uu^TH - \lambda I) &= p_A(\lambda) - p_A(\lambda) u^TH(\lambda - A)^{-1}u \\
&= \lambda^6 \left(1 + \begin{bmatrix} u_4 & -u_3 & -u_2 & u_1 & u_6 & u_5 \end{bmatrix} \begin{bmatrix} -\frac{1}{\lambda}u_1 - \frac{1}{\lambda^2}u_2 \\ -\frac{1}{\lambda}u_2 \\ -\frac{1}{\lambda}u_3 - \frac{1}{\lambda^2}u_4 \\ -\frac{1}{\lambda}u_4 \\ -\frac{1}{\lambda}u_5 - \frac{1}{\lambda^2}u_6 \\ -\frac{1}{\lambda}u_6 \end{bmatrix} \right) \\
&= \lambda^6 - \lambda^5 u_1 u_4 - \lambda^4 u_2 u_4 + \lambda^5 u_2 u_3 + \lambda^5 u_2 u_3 \\
&+ \lambda^4 u_2 u_4 - \lambda^5 u_4 u_1 - \lambda^5 u_5 u_6 - \lambda^4 u_6^2 - \lambda^5 u_5 u_6 \\
&= \lambda^4(\lambda^2 + 2\lambda(-u_1u_4 + u_2u_3 - u_5u_6) - u_6^2).
\end{aligned}$$

Thus it follows that there are generically two non-zero eigenvalues as well as two Jordan blocks of size two with eigenvalue zero.

Observe by the way that the computation above also gives insight in the first example: the characteristic polynomial of B has terms $-\lambda^4 u_2 u_4$ and $\lambda^4 u_2 u_4$ which cancel. In the first example this will also happen, and the characteristic polynomial there will be $\lambda^4 - 2\lambda^3(u_1 u_4 - u_2 u_3)$, which has a triple zero at zero as expected.

4.6 Appendix

In this appendix we first present a simple and well-known lemma which is used in the proof of part of the main theorem, Theorem 4.3.1.

Lemma 4.6.1 *For any $p \times m$ matrix X and $m \times p$ matrix Y we have*

$$\det(I_p - XY) = \det(I_m - YX).$$

Proof. The lemma is a straightforward consequence of the following identities:

$$\begin{bmatrix} I_p & X \\ Y & I_m \end{bmatrix} \begin{bmatrix} I_p & -X \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ Y & I_m - YX \end{bmatrix},$$

and

$$\begin{bmatrix} I_p & X \\ Y & I_m \end{bmatrix} \begin{bmatrix} I_p & 0 \\ -Y & I_m \end{bmatrix} = \begin{bmatrix} I_p - XY & X \\ 0 & I_m \end{bmatrix}.$$

Indeed, the first identity implies, upon taking determinants that

$$\det \begin{bmatrix} I_p & X \\ Y & I_m \end{bmatrix} = \det(I_m - YX),$$

the second identity implies, again upon taking determinants that

$$\det \begin{bmatrix} I_p & X \\ Y & I_m \end{bmatrix} = \det(I_p - XY).$$

Hence the lemma is proved. □

Our next result is useful for the considerations in Section 4.4.

Lemma 4.6.2 *Let p_1 and p_2 be real polynomials of degree n and m, respectively, with $n > m$. Assume p_1 and p_2 have no common roots. Consider for $t > 0$ the polynomial $q_t(\lambda) = p_1(\lambda) - t p_2(\lambda)$. Denote by λ_j, $j = 1, \ldots, m$ the roots of p_2. Denote by $\mu_j(t), j = 1, \ldots, n$ the roots of q_t. Then there is an ordering of the μ_j such that, as t goes to infinity we have*

$$\lim_{t\to\infty} \mu_j(t) = \lambda_j, \quad j = 1, \ldots, m.$$

Moreover, for $j > m$ we have $\lim_{t\to\infty} \mu_j(t) = \pm\infty$.

Proof. Instead of q_t we consider

$$\tilde{q_t}(\lambda) = \frac{1}{t} q_t(\lambda) = \frac{1}{t} p_1(\lambda) - p_2(\lambda).$$

Let $\varepsilon > 0$ be given, so that the closed discs $C_j = \{\lambda | \, |\lambda - \lambda_j| \leq \varepsilon\}$ do not intersect. We have to show that for t large enough, the polynomial $\tilde{q}_t$ has one root in each closed disc C_j. Now $\frac{1}{t}p_1$ converges uniformly to zero on the union of the closed discs C_j. We can then apply Hurwitz's theorem (see [67], Theorems III.11.1 and Theorem III.11.2), to obtain that there is a t_0 so that for $t > t_0$ the polynomial $\tilde{q}_t$ and hence q_t has exactly one root in each of the C_j's. Since this holds for every $\varepsilon > 0$, we have proved the first part of the lemma.

To show the statement about the remaining roots going to infinity, we argue by contradiction. So, suppose that for some complex number a we have that it is a limit of roots of q_t as t goes to infinity. Since the roots of q_t and $\tilde{q}_t$ coincide, we will consider $\tilde{q}_t$. Let $a(t)$ be a root of $\tilde{q}_t$, that is $p_1(a(t))/t - p_2(a(t)) = 0$, and suppose $a(t) \to a$. Then by continuity $p_2(a(t)) \to p_2(a)$, and $p_1(a(t)) \to p_1(a)$. In particular, it follows that $p_1(a(t))/t \to 0$ as $t \to \infty$. But then it follows that $p_2(a) = 0$. $\square$

Bibliography

[1] T. Ando. *Linear Operators on Krein Spaces*. Research Institute of Applied Electricity, Hokkaido University, Sapporo, Japan, 1979.

[2] T. Ya. Azizov. Invariant subspaces and criteria for the completeness of the system of root vectors of J-dissipative operators in the Pontrjagin Π_κ. (Russian) *Dokl. Akad. Nauk SSSR* 200, (1971), 1015–1017.

[3] T. Ya. Azizov. Dissipative operators in a Hilbert space with indefinite metric. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* 37, (1973), 639–662.

[4] T. Ya. Azizov, E. I. Iohvidov. Invariant subspaces of maximal J-dissipative operators. (Russian) *Mat. Zametki* 12, (1972), 747–754.

[5] T. Ya. Azizov, I.S. Iohvidov. *Linear operators in spaces with an indefinite metric*. John Wiley and Sons, Chicester, 1989 [Russian original 1979].

[6] H. Bart, I. Gohberg, M.A. Kaashoek, A.C.M. Ran. *Factorization of matrix and operator functions: the state space method*. OT **178**, Birkhäuser, Basel etc., 2008.

[7] H. Bart, I. Gohberg, M.A. Kaashoek, A.C.M. Ran. *A state space approach to canonical factorization with applications*. OT **200**, Birkhäuser, Basel etc., 2010.

[8] H. Baumgärtel. *Analytic Perturbation Theory for Matrices and Operators*. Birkhäuser, Basel etc., 1985.

[9] Ts. Bayasgalan. On strongly stable normal operators on Krein space. *Acta Math. Hungar*, 81, (1998), 275–278.

[10] Ts. Bayasgalan. Fundamental reducibility of normal operators on Krein space. *Stud. Sci. Math. Hung*, 35, (1999), 147–150.

[11] M.A. Beitia, I. de Hoyos and I. Zaballa. The change of the Jordan structure under one row perturbations. *Linear Algebra and its Applications*, 401, (2005), 119–134.

[12] J. Bognár. *Indefinite Inner Product Spaces*. Springer-Verlag, New York- Heidelberg, 1974.

[13] P. Brunovský. A classification of linear controllable systems. *Kybernetika (Prague)* 6 (1970), 173–188.

[14] F. Dopico, J. Moro. Low rank perturbation of Jordan structure. *SIAM J. Matrix Anal. Appl.*, 25, (2003), 495–506.

[15] J. Dieudonné. *Foundations of modern analysis.* Academic Press, Inc., New York, 1960. International edition.

[16] D.Ž. Djoković, J. Patera, P. Winternitz and H. Zassenhaus. Normal forms of elements of classical real and complex Lie and Jordan algebras. *J. Math. Phys*, 24 (1983), 1363–1373.

[17] A. Ferrante, B. C. Levy. Canonical form of symplectic matrix pencils. *Linear Algebra and its Applications*, 274, (1998), 259–300.

[18] D. T. Finkbeiner. *Introduction to Matrices and Linear Transformation.* W. H. Freeman and Co., San Francisco, California-London, 1966. Second Edition.

[19] J.H. Fourie, G. Groenewald, A.C.M. Ran. Positive real matrices in indefinite inner product spaces and invariant maximal semidefinite subspaces. *Linear Algebra and its Applications*, Vol. 424, (2007), 346-370.

[20] B.A. Francis. *A course in H_∞-control theory.* Springer Verlag, Berlin etc., 1987.

[21] F. R. Gantmacher. *The Theory of Matrices*, Volume 1. Chelsea Publishing Company, New York, N. Y., 1959.

[22] W. Greenberg, C.V.M. van der Mee, V. Protopopescu. *Boundary value problems in abstract kinetic theory.* OT **25**, Birkhäuser Verlag, Basel, 1986.

[23] I. Gohberg, S. Goldberg, M.A. Kaashoek. *Classes of Linear Operators*, Volume 1. OT **49**, Birkhäuser Verlag, 1990.

[24] I. Gohberg, P. Lancaster, L. Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser Verlag, Basel, 2005.

[25] I. Gohberg, P. Lancaster, L. Rodman. *Matrices and Indefinite Scalar Products.* OT **8**, Birkhäuser Verlag, Basel, 1983.

[26] I. Gohberg, P. Lancaster, L. Rodman. *Invariant Subspaces of Matrices with Applications.* John Wiley & Sons, Inc., New York, NY, 1986 = Classics in Appl. Math. **51**, *SIAM*, Philadelphia, PA, 2006.

[27] I. Gohberg, B. Reichstein. Classification of block-Toeplitz H-normal operators. *Linear and Multilinear Algebra* 34, (1993), 213–245.

[28] I. Gohberg, B. Reichstein. On H-unitary and block-Toeplitz H-normal operators. *Linear and Multilinear Algebra* 30, (1991), 17–48.

[29] I. Gohberg, B. Reichstein. On classification of normal matrices in an indefinite scalar product. *Integral Equations Operator Theory* 13, (1990), 364–394.

[30] O.V. Holtz, V.A. Strauss. Classification of normal operators in spaces with indefinite scalar product of rank 2. *Linear Algebra and its Applications* 241/243, (1996), 455–517.

[31] O.V. Holtz, V.A. Strauss. On classification of normal operators in real spaces with indefinite scalar product. *Linear Algebra and its Applications* 255, (1997), 113–155.

[32] O.V. Holtz. On indecomposable normal matrices in spaces with indefinite scalar product. *Linear Algebra and its Applications* 259, (1997), 155–168.

[33] L. Hörmander, A. Melin. A remark on perturbations of compact operators. *Math. Scand.* 75 (1994), 255–262.

[34] R. Horn, C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[35] I.S. Iohvidov, M.G. Krein, H. Langer. *Introduction to spectral theory of operators in spaces with indefinite metric*. Mathematical Research, Vol. 9, Akademie-Verlag, Berlin, 1982.

[36] T. Kato. *Perturbation Theory for Linear Operators*. Springer, New York, NY, 1966.

[37] M.G. Krein, H. Langer. On some mathematical principles in the linear theory of damped oscillations of continua I, II. *Integral Equations and Operator Theory* 1 (1978), 364-399; 539-566 (translation from Russian; original 1965).

[38] M. Krupnik. Changing the spectrum of an operator by perturbation. *Linear Algebra and its Applications* 167, (1992), 113–118.

[39] P. Lancaster, L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra and its Applications* 406, (2005), 1–76.

[40] P. Lancaster, L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Rev.* 47, (2005), 407–443.

[41] P. Lancaster, L. Rodman. *Algebraic Riccati equations*. Oxford University Press, 1995.

[42] H. Langer. Spectral functions of definitizable operators in Krein spaces. *Proc. Graduate School "Functional Analysis", Dubrovnik 1981. Lecture Notes in Math.* 948, Springer Verlag, Berlin, 1982, 1-46.

[43] P. Lancaster, M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, 1985.

[44] A. J. Laub, K. R. Meyer. Canonical Forms for Symplectic and Hamiltonian Matrices. *Celestial Mech.* 9, (1974), 213–238.

[45] V.B. Lidskii. To perturbation theory of non-selfadjoint operators. *Zh. Vychisl. Mat. i Mat. Fiz.* (*U.S.S.R. Comput. Math. and Math. Phys.*) 6, (1966), 52–60.

[46] W.-W. Lin, V. Mehrmann, H. Xu. Canonical Forms for Hamiltonian and Symplectic Matrices and Pencils. *Linear Algebra and its Applications* 301–303, (1999), 469–533.

[47] C.V.M. van der Mee, A.C.M. Ran, L. Rodman. Stability of stationary transport equations with accretive collision operators. *Journal of Functional Analysis*, 174 (2000), 478-512.

[48] Chr. Mehl. On classification of polynomially normal matrices in indefinite inner product spaces. *Electron. J. Linear Algebra* 15, (2006), 50–83.

[49] Chr. Mehl. Essential decomposition of normal matrices in real indefinite inner product spaces. *Electron. J. Linear Algebra* 15, (2006), 84–106.

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

[51] Chr. Mehl, V. Mehrmann, H. Xu. Canonical Forms for Doubly Structured Matrices and Pencils. *Electron. J. Linear Algebra* 7, (2000), 112–151.

[52] Chr. Mehl, V. Mehrmann, A.C.M. Ran, L. Rodman. Perturbation analysis of Lagrangian invariant subspaces of symplectic matrices. *Linear Multilinear Algebra* 57, (2009), 141–184.

[53] Chr. Mehl, V. Mehrmann, A.C.M. Ran, L. Rodman. Eigenvalue perturbation theory of classes of structured matrices under generic structured rank one perturbations. *Linear Algebra and its Applications* 435, (2011), 687–716.

[54] Chr. Mehl, V. Mehrmann, A.C.M. Ran, L. Rodman. Perturbation theory of selfadjoint matrices and sign characteristics under generic structured rank one perturbations. *Linear Algebra and its Applications* 436, (2012), 4027-4042.

[55] Chr. Mehl, V. Mehrmann, A.C.M. Ran, L. Rodman. Jordan forms of real and complex matrices under generic rank one perturbations. To appear in *Operators and Matrices*.

[56] C. Mehl, L. Rodman. Classes of normal matrices in indefinite inner products. *Linear Algebra and its Applications* 336, (2001), 71-98.

[57] J. Moro, J.V. Burke, M.L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.* 18, (1997), 793–817.

[58] L.S. Pontryagin. Hermitian operators in spaces with indefinite metric. (Russian) *Izvestija Akad. Nauk SSSR Ser. Mat.* 8, (1944), 243–280.

[59] A.C.M. Ran, L. Rodman. Stability of invariant maximal semidefinite subspaces I. *Linear Algebra and its Applications* 62, (1984), 51–86.

[60] A.C.M. Ran, L. Rodman. Stability of invariant maximal semidefinite subspaces II. Applications: Self-adjoint rational matrix functions, algebraic Riccati equations. *Linear Algebra and its Applications* 63, (1984), 133–173.

[61] A.C.M. Ran, L. Rodman. Stability of invariant Lagrangian subspaces I. OT **32**, Birkhäuser Verlag, Basel, (1988), 181–218.

[62] A.C.M. Ran, L. Rodman. Stability of invariant Lagrangian subspaces II. OT **40**, Birkhäuser Verlag, Basel, (1989), 391–425.

[63] A.C.M. Ran, L. Rodman, D. Temme. Stability of pseudo-spectral factorizations. In: *Operator Theory and Analysis, The M.A. Kaashoek Anniversary Volume*. OT **122**, Birkhäuser Verlag, Basel, (2001), 359–383.

[64] A.C.M. Ran, D. Temme. Invariant semidefinite subspaces of dissipative matrices in an indefinite inner product space, existence, construction and uniqueness. *Linear Algebra and its Applications* 212/213, (1994), 169–214.

[65] A.C.M. Ran, M. Wojtylak. Eigenvalues of rank one perturbations of unstructured matrices. *Linear Algebra and its Applications* 437, (2012), 589–600.

[66] L. Rodman. Similarity vs unitary similarity: Complex and real indefinite inner products. *Linear Algebra and its Applications* 416, (2006), 945–1009.

[67] S. Saks, A. Zygmund. *Analytic Functions, 3rd Edition.* Elsevier Publishing Company, 1971.

[68] S.V. Savchenko. On a generic change in the spectral properties under perturbation by an operator of rank one. (Russian) *Mat. Zametki* 74, (2003), 590–602; translation in *Math. Notes* 74, (2003), 557–568.

[69] S.V. Savchenko. On the change in the spectral properties of a matrix under a perturbation of a sufficiently low rank. (Russian) *Funktsional. Anal. i Prilozhen.* 38 (2004), 85–88; translation in *Funct. Anal. Appl.* 38, (2004), 69–71.

[70] S.V. Savchenko. Laurent expansion for the determinant of the matrix of scalar resolvents. *Mathematicheskii Sbornik* 196, (2005), 121–144; translation in *Sbornik Mathematics* 196, (2005), 743–764.

[71] A.A. Shkalikov. Selection principles and properties of some proper and adjoined elements of operator pencils. *Moscow Univ. Math. Bull.* **43**, (1988), 16–25.

[72] A.A. Shkalikov. On invariant subspaces of dissipative operators in a space with indefinite metric. (Russian) *Tr. Mat. Inst. Steklova* 248, (2005), 294–303. (English translation in *Proc. of Steklov Inst.* to appear.)

[73] A.A. Shkalikov. Invariant subspaces of dissipative operators in Krein space and Sobolev problem on a rotating top. In: *Functional analysis and its applications*, North-Holland Math. Stud., 197, Elsevier, Amsterdam (2004), 273–286.

[74] G.E. Shilov. *An Introduction to the Theory of Linear Spaces.* Dover Publications, New York, 1974. (Translation from Russian.)

[75] L. Smith. *Linear Algebra*. Springer-Verlag, New York, 1984.

[76] S.L. Sobolev. The motion of a symmetric top containing a cavity filled with a liquid (Russian). *Z. Prikl. Meh. i Tehn. Fiz* 3, (1960), 20–55.

[77] D. Temme. *Dissipative Operators in Indefinite Scalar Product Spaces.* Ph.D. thesis, Vrije Universiteit, Amsterdam, 1996.

[78] R.C. Thompson. Pencils of Complex and Real Symmetric and Skew Matrices. *Linear Algebra and its Applications* 14, (1991), 323–371

[79] R.C. Thompson. Invariant factors under rank one perturbations. *Canad. J. Math.* 32, (1980), 240–245.

[80] M.I. Vishik, L.A. Lyusternik. Solutions of some perturbation problems in the case of matrices and self-adjoint and non-self-adjoint differential equations. *Uspehi Mat. Nauk* (*Russian Math. Surveys*) 15, (1960), 3–80.

[81] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford Univ. Press, Oxford, 1965.

[82] C. Xiaoman, H. Chaocheng. Normal operators in Π_κ-space. *Northeast. Math. J*, 1, (1985), 247–252.

Index

A-invariant neutral subspace, 24
A-invariant nonnegative subspace, 24
A-invariant nonpositive subspace, 24
H-dissipative matrices, 19
H-expansive, 46
H-orthogonal, 83, 85
H-positive real, 19, 96
H-unitary, 84
$\mathbb{A}$-stable, 38
$\mathbb{D}$-stable, 38

algebraic multiplicity, 110, 114
asymptotically, 115

Cauchy-Schwarz inequality, 97
Cayley transform, 46
characteristic matrix, 32
characteristic polynomial, 109, 111

degenerate, 19

even numerical range, 33
even numerical range condition, 36

generic set, 96
geometric multiplicity, 99
Gram matrix, 21

Hankel structure, 22
Hurwitz's theorem, 123

indefinite inner product, 19

Jordan basis, 25
Jordan chain, 19, 108, 111

length of Jordan chain, 19

minimal polynomial, 112

neutral subspace, 22
nondegenerate, 19
numerical range, 32

odd numerical range condition, 32

partial multiplicities, 99
perturbation, 96
perturbed matrix, 100
positive semidefinite, 97

rank-k matrices, 97
rank-one H-positive real matrix, 100
rank-one matrix, 96
rank-one perturbation, 96
rank-one perturbation matrix, 111

spectral projection, 112
Sylvester resultant matrix, 114

unstructured rank-one perturbations, 100