HOLOMORPHIC AND MEROMORPHIC EIGENVALUE PROBLEMS: THEORY AND NUMERICS

Ph.D. Thesis

JIBRAIL ALI

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI - 781039, INDIA

August 2024

HOLOMORPHIC AND MEROMORPHIC EIGENVALUE PROBLEMS: THEORY AND NUMERICS

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Jibrail Ali

(Roll No. 146123009)

Under the Supervision of

Prof. Rafikul Alam

to the

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI - 781039, ASSAM, INDIA

August 2024

DECLARATION

I do hereby declare that the work contained in this thesis entitled **"Holomorphic and Meromorphic Eigenvalue Problems: Theory and Numerics"** has been done by me, a student in the Department of Mathematics, Indian Institute of Technology Guwahati under the guidance of **Prof. Rafikul Alam** for the award of the degree of Doctor of Philosophy and that this work has not been submitted elsewhere for a degree.

August 2024

Jibrail Ali
(Roll No. 146123009)
Department of Mathematics
Indian Institute of Technology Guwahati

CERTIFICATE

It is certified that the work contained in this thesis entitled **"Holomorphic and Meromorphic Eigenvalue Problems: Theory and Numerics"** by **Jibrail Ali,** a student in the Department of Mathematics, Indian Institute of Technology Guwahati, for the award of the degree of Doctor of Philosophy has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

August 2024

Dr. Rafikul Alam
Professor
Department of Mathematics
Indian Institute of Technology Guwahati

Dedicated To my wife

Razia Begum

and my sons
Muhammad Shahid and Muhammad Hujaifa

ACKNOWLEDGMENT

First and foremost, I want to convey my deepest and most heartfelt gratitude to my thesis supervisor, **Prof. Rafikul Alam** , for allowing me to work under his guidance for my PhD. This thesis is a collaborative effort in which he made a remarkable contribution. I sincerely appreciate his unwavering dedication, hard work, and patience in supporting me. His expertise and straightforward research approach have been a tremendous source of motivation. I am genuinely thankful for his invaluable guidance, consistent encouragement, and careful attention throughout this process. His readiness to address my numerous questions and offer insightful feedback made my PhD experience both productive and enriching. I am profoundly grateful for his steadfast support and mentorship.

I would also like to extend my sincere thanks to my doctoral committee members, Prof. Rajen Kumar Sinha, Prof. Shreemayee Bora, and Dr. Anjan Kumar Chakrabarty, for their valuable comments, suggestions, and concern for my research progress. My gratitude also extends to all the faculty members of the Department of Mathematics, IIT Guwahati, for their support throughout my thesis work.

I am thankful to the Indian Institute of Technology Guwahati for the facilities provided during my research work and the Ministry of Human Resource Development, Govt. of India, for the financial assistance to pursue my PhD at IIT Guwahati. I also appreciate the support from all the staff members of the Department of Mathematics, IIT Guwahati, in both administrative and technical matters.

I am incredibly grateful to my employer, the Indian Institute of Technology (BHU) Varanasi, for encouraging me to pursue higher studies and providing Extraordinary Leave (EOL). Special thanks to Prof. Rajeev Sangal, the first Director of IIT (BHU), for his special attention and blessings. I acknowledge the Department of Chemistry, IIT (BHU), for their cooperation and support throughout my PhD. My heartfelt thanks go to my colleagues Jagadish Kumar and Sundeep, and faculty members Prof. Jeykumar Kandasamy, Dr. Pandeshwar Makam, Dr. Rehan and Prof. D. Tiwary for their encouragement and guidance. I am also highly thankful to Dr. Shaivya Singh for sharing ideas and constantly motivating me.

Special acknowledgment goes to Ranjan Das for his friendship. I also want to thank Jogen Datta, Nandita, and Dipankar for making my time at IIT Guwahati enjoyable and memorable. Special thanks to Dr. Ranjan for his help with LaTeX and MATLAB, and to Md. Sohidul Alam for his contributions and suggestions, including assistance with the final proofreading of the thesis. I appreciate the companionship of friends like Kannan R, Sazee, Rafi Dani, Nadim Atif, and Naseem Akhtar. Additionally, I enjoyed the company of Prof. Sheik Rafi Ahmed from the Department of EEE at IIT Guwahati.

Words cannot fully express my gratitude towards my parents, Peer Mohammad and Madina Begum, my brother Israel Ali, and my wife Razia Begum for their unconditional love, care, support, and belief in me. I am especially grateful to my wife and my sons, Muhammad Shahid and Muhammad Hujaifa, for their love and care. Conversations with my hearts, Muhammad Shahid and Muhammad Hujaifa, always brought me joy and amusement.

Finally, I wish to express my utmost gratitude to my Creator, the Sovereign, the Knowledge Provider, the Sustainer, the Protector, and the Provider of all my needs. I sincerely thank YOU for guiding me through this journey. YOUR support gave me the resilience to overcome challenges, and YOUR endless blessings and guidance empowered me to complete this thesis. I am eternally grateful for YOUR boundless grace and mercy.

Above all, once again, I thank YOU for everything!

August 2024

With Regards,

Jibrail Ali

ABSTRACT

Numerics of nonlinear eigenvalue problems is a challenging subject and an active area of research. The thesis addresses three broad problems in the area of theory and numerics of nonlinear eigenvalue problems. The first problem concerns with spectral analysis of holomorphic and meromorphic eigenvalue problems. The second problem concerns with perturbation theory of holomorphic eigenvalue problems as well as realization and linearization of meromorphic matrices. The third problem concerns with numerical solution of holomorphic and meromorphic eigenvalue problems.

Canonical forms such as Smith forms of matrix polynomials and Smith-McMillan forms of rational matrices are important tools for spectral analysis of matrix polynomials and rational matrices. We present, among other things, global Smith forms for holomorphic matrices and global Smith-McMillan forms for meromorphic matrices, which are akin to Smith forms and Smith-McMillan forms of matrix polynomials and rational matrices, for spectral analysis of holomorphic and meromorphic matrices . Also, we present a detailed system theoretic analysis (e.g. matrix fraction descriptions (MFDs), realizations, and analytic system matrices) of meromorphic matrices and establish relationships between canonical forms of meromorphic matrices, their MFDs and analytic system matrices.

Linearization is an important method for solving and analyzing polynomial and rational eigenvalue problems. We utilize linearization of holomorphic operator-valued functions and develop perturbation theory for holomorphic eigenvalue problems. We derive perturbation bounds for (discrete) eigenvalues and eigenvectors of holomorphic operator-valued functions. We analyze local minimal realizations of meromorphic matrices and compute such realizations from Markov parameters. We utilize local minimal realization of a meromorphic matrix to define and construct a linearization. The linearizations of meromorphic matrix-valued functions can be utilized for developing perturbation theory for meromorphic eigenvalue problems.

Finally, we show that numerical methods for solving holomorphic eigenvalue problems can be developed from pole finding methods for meromorphic matrices. Utilizing this point of view, we provide a unified framework for contour integration based methods for solving nonlinear eigenvalue problems. We propose and solve spectral recovery problems, which are actually moment problems of certain kind, for linear and nonlinear eigenvalue problems in finite and infinite dimension.

Contents

List of Symbols

$\mathbb{N}$	the set of the natural numbers
$\mathbb{Z}$	the set of the integers
$\mathbb{R}$	the field of the real numbers
$\mathbb{C}$	the field of the complex numbers
$\mathbb{C}^{m\times n}$	the space of complex matrices of size $m \times n$
$\mathbb{C}[z]$	the ring of polynomials
$\mathbb{C}[z]^{m\times n}$	the set of $m \times n$ matrix polynomials
$\mathbb{C}(z)$	the field of rational functions
$\mathbb{C}(z)^{m\times n}$	the vector space of $m \times n$ rational matrices
I_n	the identity matrix of size n
e_k	k-th column of the identity matrix I_n
$A^\top$	the transpose of real A
A^*	complex conjugate transpose of A

1 Introduction

Nonlinear eigenvalue problems arise in many applications and their numerical treatment is an active area of research; see [10, 11, 42, 43, 44, 29, 45] and the references therein. For example, consider the delay-differential equation (DDE) [45]

$$\frac{d}{dt}x(t) = A_0 x(t) + \sum_{i=1}^{m} A_i x(t - \tau_i),$$

where $x(t) \in \mathbb{C}^n$ is a vector of state variable at time $t \in \mathbb{R}$, $A_1, \ldots, A_m$ are matrices in $\mathbb{C}^{n\times n}$ and $0 < \tau_1 < \tau_2 < \cdots < \tau_m$ are the time-delays. Now seeking a solution of the form $x(t) = ve^{\lambda t}$ with $v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ yields the nonlinear eigenvalue problem (NEP)

$$\mathbf{T}(\lambda)v := (\lambda I - A_0 - \sum_{i=1}^{m} A_i e^{-\lambda \tau_i})v = 0.$$

Nonlinear eigenvalue problems involving rational matrices (i.e., matrices whose entries are rational functions) also arise in applications. For example, the rational eigenvalue problem (REP)

$$G(\lambda)v := -Av + \lambda Bv + \lambda^2 \sum_{j=1}^{k} \frac{1}{w_j - \lambda} C_j v = 0$$

arises when a generalized linear eigenvalue problem condensed exactly [43]. The state-space analysis of linear time-invariant systems also leads to rational eigenvalue problems in which the rational matrices arise as the transfer functions, see [32, 15, 56].

Nonlinear eigenvalue problems involving meromorphic matrices (i.e., matrices whose entries are meromorphic functions) also arise in applications such as in the study of controllability of delay-differential systems (DDS) [18, 19, 59] given by

$$\begin{aligned} \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} &= A_0\mathbf{x}(t) + \sum_{j=1}^{N_1} A_j\mathbf{x}(t-\tau_j) + \sum_{j=1}^{N_2} B_j\mathbf{u}(t-t_j), \\ \mathbf{y}(t) &= \sum_{j=1}^{N_3} C_j\mathbf{x}(t-s_j) + \sum_{j=1}^{N_4} D_j\mathbf{u}(t-h_j). \end{aligned}$$

Here $\mathbf{x}(t) \in \mathbb{C}^r$ is a vector of state variables and $\mathbf{u}(t) \in \mathbb{C}^n$ is a vector of control variables. Further, $0 < \tau_1 < \cdots < \tau_{N_1}$, $0 \le t_1 < \cdots < t_{N_2}$, $0 \le s_1 < \cdots < s_{N_3}$ and $0 \le h_1 < \cdots < h_{N_4}$ are delay parameters. Furthermore, $A_0, \ldots, A_{N_1}$ are $r \times r$ constant matrices, $B_1, \ldots, B_{N_2}$ are $r \times n$ constant matrices, $C_1, \ldots, C_{N_3}$ are $m \times r$ constant matrices and $D_1, \ldots, D_{N_4}$ are $m \times n$ constant matrices.

For the ansatz $\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} := \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} e^{\lambda t}$ with $\mathbf{x} \in \mathbb{C}^{r\times r}$ and $\mathbf{u} \in \mathbb{C}^{n\times n}$, the DDS yields

$$\begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} e^{\lambda t} = \begin{bmatrix} 0 \\ \mathbf{y}(t) \end{bmatrix} \tag{1.1}$$

where $A(\lambda) := \lambda I_r - A_0 - \sum_{j=1}^{N_1} A_j e^{-\lambda\tau_j}$, $B(\lambda) := \sum_{j=1}^{N_2} B_j e^{-\lambda t_j}$, $C(\lambda) := \sum_{j=1}^{N_3} C_j e^{-\lambda s_j}$ and $D(\lambda) := \sum_{j=1}^{N_4} D_j e^{-\lambda h_j}$. The holomorphic matrix (i.e., matrix whose entries are holomorphic functions)

$$\mathbf{S}(\lambda) := \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix}$$

is called the system matrix of the DDS and the meromorphic matrix

$$\mathbf{M}(\lambda) := D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$$

is called the transfer function of the DDS. Eliminating $\mathbf{x}e^{\lambda t}$ in (1.1, we have

$$\mathbf{y}(t) = [D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)]\mathbf{u}e^{\lambda t} = \mathbf{M}(\lambda)\mathbf{u}e^{\lambda t}.$$

Observe that $\mathbf{y}(t) = 0$ whenever $\mathbf{M}(\lambda)\mathbf{u} = 0$, that is, when $(\lambda, \mathbf{u})$ is an eigenpair of $\mathbf{M}(z)$. Alternatively, it follows from (1.1) that $\mathbf{y}(t) = 0$ whenever $\mathbf{S}(\lambda) \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = 0$. This shows that zero output of the DDS can be analyzed via eigenvalues and eigenvectors of the system matrix $\mathbf{S}(\lambda)$ as well as via eigenvalues and eigenvectors of the transfer function $\mathbf{M}(\lambda)$.

Motivated by above examples, we undertake a detailed analysis of theory and numerics of holomorphic and meromorphic eigenvalue problems. We always assume that $\Omega \subset \mathbb{C}$ is open and connected. Let $\mathbf{T} : \Omega \longrightarrow \mathbb{C}^{m\times n}$ be holomorphic. We consider the holomorphic eigenvalue problem

$$\mathbf{T}(\lambda)v = 0 \tag{1.2}$$

and undertake a global spectral analysis of $\mathbf{T}(z)$. We also discuss spectral perturbation theory and develop numerical methods for solving (1.2) when $\mathbf{T}(z)$ is regular. Next, we consider a meromorphic matrix-valued function $\mathbf{M} : \Omega \longrightarrow \mathbb{C}^{m\times n}$ and the meromorphic eigenvalue problem

$$\mathbf{M}(\lambda)v = 0. \tag{1.3}$$

We undertake a global spectral analysis of $\mathbf{M}(z)$ via transfer function realizations and system matrices. We also introduce linearization of $\mathbf{M}(z)$ via local minimal realization of $\mathbf{M}(z)$ which can be gainfully utilized for perturbation analysis of the meromorphic eigenvalue problem (1.3).

The thesis addresses three broad problems in the area of theory and numerics of nonlinear eigenvalue problems. The first problem concerns with spectral analysis of holomorphic and meromorphic eigenvalue problems. The second problem concerns with perturbation theory of holomorphic eigenvalue problems as well as realization and linearization of meromorphic matrices. The third problem concerns with numerical solution of holomorphic eigenvalue problems.

Local spectral theory for regular holomorphic matrix-valued functions has been developed in [24]. Also, local spectral analysis of a regular holomorphic operator-valued function has been investigated, for example, in [23, 27, 28, 44]. Canonical forms are important tools for spectral analysis. The Smith canonical form [25, 32, 50, 56] of an $m \times n$ matrix polynomial $P(z)$ is a well known canonical form which provides a complete information about the spectral data of $P(z)$. Similarly, the Smith-McMillan form [32, 50, 56, 9] of an $m \times n$ rational matrix $G(z)$ is a well known canonical form which provides a complete information about the spectral data of $G(z)$. A local version of the Smith canonical form of a holomorphic matrix-valued function $\mathbf{T} : \Omega \longrightarrow \mathbb{C}^{m\times n}$ has been studied in the literature when $\mathbf{T}(z)$ is regular; see [24, 9, 35, 44]. Likewise, a local version of the Smith-McMillan form of a meromorphic matrix-valued function $\mathbf{M} : \Omega \longrightarrow \mathbb{C}^{m\times n}$ has been studied in the literature when $\mathbf{M}(z)$ is regular; see [9, 26, 35]. Further, a global canonical form for $\mathbf{M}(z)$ has been presented in [38] where it is shown that $\mathbf{M}(z)$ is "equivalent" to a diagonal matrix $\mathrm{diag}(f_1(z), \ldots, f_r(z), 0, \ldots, 0)$. The functions $f_1(z), \ldots, f_r(z)$ are meromorphic on Ω and have the property that $f_i(z)$ divides $f_{i+1}(z)$ for $i = 1 : r-1$.

Canonical forms such as Smith forms of matrix polynomials and Smith-McMillan form of rational matrices are important tools for spectral analysis of matrix polynomials and rational matrices. Although, local canonical forms such as local Smith forms of holomorphic/meromorphic matrices are well known in the literature, there has been relatively little research on global canonical forms of holomorphic and meromorphic matrices. We will fill this gap. We utilize Weierstrass product and Weierstrass factorization of holomorphic functions and present, among other things, global Smith forms for holomorphic matrices and global Smith-McMillan forms for meromorphic matrices which are akin to Smith forms and Smith-McMillan forms of matrix polynomials and rational matrices.

Linear systems theoretic analysis (e.g., matrix fraction descriptions (MFDs), realization theory, and system matrices) of rational matrices are well known in the literature [32, 56, 50] in which Smith form of matrix polynomials and Smith-McMillan form of rational matrices play an important role. We utilize global Smith forms of holomorphic matrices and global Smith-McMillan forms of meromorphic matrices to develop global system theoretic analysis (e.g. MFDs, realization theory, and analytic system matrices) of meromorphic matrices and establish relationships between canonical forms of meromorphic matrices, their MFDs and analytic system matrices. These results are akin to corresponding results for rational matrices.

Linearization is a powerful method widely used for analyzing nonlinear eigenvalue problems such as polynomial eigenvalue problems and rational eigenvalue problems. Consider an $n \times n$ polynomial eigenvalue problem (PEP)

$$P(\lambda)v := (A_0 + \lambda A_1 + \cdots + \lambda^m A_m)v = 0.$$

For theory and computation, the PEP is transformed to an equivalent $mn \times mn$ linear eigenvalue problem of the form

$$(\mathcal{A} + \lambda\mathcal{B})\mathbf{v} = 0$$

by means of linearization. The matrix pencil $L(\lambda) := \mathcal{A} + \lambda\mathcal{B}$ is a called a linearization of $P(\lambda)$ if there exist $mn \times mn$ unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ such that

$$U(\lambda)L(\lambda)V(\lambda) = \mathrm{diag}(I_{(m-1)n},\ P(\lambda)).$$

Linearization of matrix polynomials has been studied extensively over the years, see [25, 37, 41, 57, 16, 49, 13] and the references therein.

For spectral analysis of holomorphic eigenvalue problems, the concept of linearization has been extended to holomorphic operator-valued functions. Let $\mathbf{A} : \Omega \longrightarrow L(X)$ be holomorphic, where X is a complex Banach space and $L(X)$ is the Banach space of all bounded linear operators on X. Define $\mathbb{T} : C(\Gamma, X) \longrightarrow C(\Gamma, X)$ by

$$(\mathbb{T}f)(z) := zf(z) - \frac{1}{2\pi i} \int_{\partial\Omega} (I - \mathbf{A}(w))f(w)dw, \tag{1.4}$$

where $C(\Gamma, X)$ is the Banach space of continuous functions $f : \Gamma \longrightarrow X$ equipped with the supremum norm. Then it is shown in [27, 23] that $\mathbb{T}$ is a bounded linear operator and there exists a Banach space Z such that $\mathbf{A}(z) \oplus I_Z$ is "analytically equivalent" to $\mathbb{T} - zI$ on $\mathcal{O} := \mathrm{Int}(\Gamma)$. Consequently, $\mathbb{T}$ and $\mathbf{A}(z)$ have the same spectra in $\mathcal{O}$. We refer to $\mathbb{T}$ as the Gohberg-Kaashoek-Lay linearization of $\mathbf{A}(z)$ or in short GKL-linearization of $\mathbf{A}(z)$ on $\mathcal{O}$. See [44] for local spectral analysis of Fredholm operator-valued functions.

Spectral perturbation theory for linear operators is a well-established classical field [39, 33]. By contrast, spectral perturbation theory for nonlinear eigenvalue problems has not received much attention. This thesis endeavors to address this problem. We consider an operator-valued function $\mathbf{V} : \Omega \longrightarrow L(X)$ and analyze the perturbed nonlinear eigenvalue problem

$$(\mathbf{A}(\lambda) + \mathbf{V}(\lambda))u = 0.$$

More precisely, we analyze the discrete spectra of the one-parameter family of holomorphic operator-valued function $W(z, t) := \mathbf{A}(z) + t\mathbf{V}(z)$ when t varies in $\mathbb{C}$. We prove, among other things, that if

$$\max_{z\in\Gamma} r_\sigma(\mathbf{V}(z)\mathbf{A}(z)^{-1}) < 1,$$

where $r_\sigma(T)$ denotes the spectral radius of an operator T, then the operator analogue of Rouche's theorem holds, that is, both $\mathbf{A}(z)$ and $\mathbf{A}(z) + \mathbf{V}(z)$ are Fredholm operators of index zero for all $z \in \mathcal{O} := \mathrm{Int}(\Gamma)$ and they have the same number (counting multiplicities) of discrete eigenvalues in $\mathcal{O}$. Further, if μ is a discrete eigenvalue of $\mathbf{A}(z)$ with algebraic multiplicity ℓ, then we determine a disc $\partial_\Gamma \subset \mathbb{C}$ such that $\mathcal{O}$ contains exactly ℓ discrete eigenvalues (counting multiplicities) $\mu_1(t), \ldots, \mu_\ell(t)$ of $W(\lambda, t)$ for all $t \in \partial_\Gamma$. Moreover, we show that the map $t \mapsto \mu_{\mathrm{av}}(t) := (\mu_1(t) + \cdots + \mu_\ell(t))/\ell$ is holomorphic on ∂_Γ. Finally, if ν is the ascent of μ then we show that there exist positive real numbers

δ, τ and real constants α, β (independent of t and $\mathbf{V}(z)$) such that

$$\|t\mathbf{V}\| < \delta \implies |\mu_j(t) - \mu|^{\nu} \leq \alpha\|t\mathbf{V}\| \quad \text{for} \quad j = 1, 2, \ldots, \ell,$$
$$\|t\mathbf{V}\| < \tau \implies |\mu_{\mathrm{av}}(t) - \mu| \leq \beta\|t\mathbf{V}\|.$$

Realization of rational matrices is a classical topic in Linear Systems Theory [32, 50, 56, 15]. It is well-known that an $m \times n$ a rational matrix $G(z)$ admits a minimal realization of the form

$$G(z) = P(z) + C(zI_r - A)^{-1}B,$$

where $P(z)$ is a matrix polynomial and $(C, A, B) \in \mathbb{C}^{n\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$ is such that (C, A) is an observable pair and (A, B) is a controllable pair. It is also well-known that poles of $G(z)$ are the eigenvalues of A and the zeros of $G(z)$ are the eigenvalues of the Rosenbrock system matrix [32, 50, 56]

$$S(z) := \left[\begin{array}{c|c} P(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

With a view to solving rational eigenvalue problems, linearization of a rational matrix is introduced in [2, 53] and further developed in [3, 14, 4]. It is shown that zeros of $G(z)$ can be analyzed by a system matrix of the form

$$\mathbb{L}(z) := \left[\begin{array}{c|c} \mathcal{X} - z\mathcal{Y} & \mathcal{C} \\ \hline \mathcal{B} & A - zI_r \end{array}\right]$$

where $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{B}$, and $\mathcal{C}$ are appropriate matrices. The pencil $\mathbb{L}(z)$ is called a linearization of $G(z)$. Thus spectral analysis of rational matrices can be undertaken via realization followed by a linearization. We extend this template to meromorphic matrices.

Let $\mathbf{M} : \Omega \longrightarrow \mathbb{C}^{m\times n}$ be meromorphic. We show that $\mathbf{M}(z)$ admits a local minimal realization on a compact set $\mathrm{K} \subset \Omega$ of the form

$$\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$$

where $\mathbf{H} : \mathrm{K} \longrightarrow \mathbb{C}^{m\times n}$ is holomorphic and $(C, A, B) \in \mathbb{C}^{n\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$ is such that (C, A) is an observable pair and (A, B) is a controllable pair. We show that the poles of $\mathbf{M}(z)$ in K are the eigenvalues of A and the zeros of $\mathbf{M}(z)$ in K are the eigenvalues of the analytic system matrix

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

We show that the matrices (C, A, B) can be determined from the Markov parameters and discuss two algorithms for computing (C, A, B). Thus, as a byproduct, we obtain a numerical method for solving a regular NEP $\mathbf{T}(\lambda)v = 0$ on a compact set by considering a local minimal realization of the meromorphic matrix $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$.

For perturbation analysis of a regular meromorphic eigenvalue problem $\mathbf{M}(\lambda)v = 0$, we introduce a linearization of $\mathbf{M}(z)$ which extends the linearization in (1.4) to the case of regular meromorphic matrices and is akin to a linearization of rational matrices. Indeed, we construct a linearization of $\mathbf{M}(z)$ of the form

$$\mathbb{L}(z) := \left[\begin{array}{c|c} zI - \mathbb{T} & \mathcal{C} \\ \hline \mathcal{B} & A - zI_r \end{array}\right],$$

where $\mathbb{T}$ is the GKL-linearization of $\mathbf{H}(z)$, $\mathcal{C}$ and $\mathcal{B}$ are appropriate operators acting on Banach spaces, and show that $\mathbf{M}(z)$ and $\mathbb{L}(z)$ have the same spectrum in K. We also describe recovery of eigenvectors of $\mathbf{M}(z)$ from those of $\mathbb{L}(z)$.

Solving nonlinear eigenvalue problems is a challenging task [42, 43]. Numerics of nonlinear eigenvalue problem is an active area of research and has received a lot of attention in recent years [10, 11, 29, 60]. The survey article [29] provides an overview of theory as well as various numerical methods for solving NEPs. Numerical methods for solving the NEP

$$\mathbf{T}(\lambda)v = 0$$

can be classified into three broad categories: (a) Newton iteration based methods (b) methods based on approximating $\mathbf{T}(z)$ (approximation by matrix polynomials, rational matrices, or projection onto lower-dimensional subspaces) and (c) methods based on contour integration. In this thesis, we focus on contour integration based method which has been studied extensively in recent years; see [60, 6, 5, 11, 12, 21, 55] and the references therein.

Observe that λ is an eigenvalue of $\mathbf{T}(z) \Longleftrightarrow \lambda$ is a pole of $\mathbf{T}(z)^{-1}$. Hence computing eigenvalues of $\mathbf{T}(z)$ is equivalent to computing poles of $\mathbf{T}(z)^{-1}$. The contour integration based methods directly or indirectly utilize this fact by invoking Keldysh theorem [34]. Hence contour integration based methods can be thought of as offshoots of local minimal realization of $\mathbf{T}(z)^{-1}$. Naturally these methods utilize Markov parameters

$$A_k = \frac{1}{2\pi i}\int_{\Gamma} z^k \mathbf{T}(z)^{-1} dz \text{ for } k \in \mathbb{Z}_+$$

or appropriate restrictions of A_k such as $A_k U$ for an arbitrarily chosen full column rank matrix $U \in \mathbb{C}^{n\times r}$ [11, 12] or $u^\top A_k v$ for arbitrarily chosen vectors u and v [60, 6, 5].

We investigate an alternative method based on tracial moments for determining eigenvalues and eigenvectors, which we refer to as spectral recovery problem or moment problem for eigenvalues. We deal with linear and nonlinear eigenvalue problems separately. The spectral recovery problem for a bounded linear operator can be stated as follows. Let $A \in L(X)$. Let $\sigma(A)$, $\sigma_d(A)$ and $\rho(A)$, respectively, be the spectrum, discrete spectrum and resolvent set of A. Suppose that $\Gamma \subset \rho(A)$ and $\sigma_0 := \sigma(A) \cap \operatorname{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. Then the spectral projection of A corresponding to the eigenvalues σ_0 is given by

$$P := \frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} dz.$$

We consider the following moment problem for the eigenvalues $\lambda_1, \ldots, \lambda_\ell$.

Problem-A: *For $k \in \mathbb{Z}_+$, consider the tracial moment*

$$s_k := \operatorname{Tr}\left(\frac{1}{2\pi i}\int_\Gamma z^k (zI - A)^{-1} dz\right).$$

Given the moments $s_0, \ldots, s_p$, for an appropriate $p \in \mathbb{N}$, determine the ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A in $\operatorname{Int}(\Gamma)$ and their algebraic multiplicities.

We show that exactly 2ℓ moments, namely, $s_0, \ldots, s_{2\ell-1}$ are needed in order to determine the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ and their algebraic multiplicities. We show that $\lambda_1, \ldots, \lambda_\ell$ are simple eigenvalues of an $\ell \times \ell$ Hankel matrix pencil $\widehat{H}_\ell - \lambda H_\ell$, where H_ℓ is nonsingular. Next, we consider recovery of the spectral projection P_j of A corresponding to the eigenvalue λ_j for $j = 1 : \ell$ from the operator moments.

Problem-B: *For $k \in \mathbb{Z}_+$, consider the operator moment*

$$S_k := \frac{1}{2\pi i}\int_\Gamma z^k (zI - A)^{-1} dz.$$

Given the moments $S_0, \ldots, S_p$, for an appropriate $p \in \mathbb{N}$, determine the spectral projections $P_1, \ldots, P_\ell$ of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively.

We show that there exists $N \in \mathbb{N}$ such that $P_1, \ldots, P_\ell$ can be recovered from N operator moments $S_0, \ldots, S_N$. However, it is not clear how to determine N. In contrast, we show that exactly ℓ operator moments, namely, $S_0, \ldots, S_{\ell-1}$ are needed in order to recover the spectral projections $P_1, \ldots, P_\ell$ when the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ are semisimple. Indeed, let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ and let the eigenvector v_j be given by $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$, $j = 1 : \ell$. Then for $j = 1 : \ell$, we show that the spectral projection P_j is given by

$$P_j = \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j = \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI - A)^{-1} dz/\alpha_j,$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ and $\alpha_j := q_{v_j}(\lambda_j)$.

Next, we consider the spectral recovery problems for holomorphic operator-valued functions by appropriately modifying the moments. Let $\mathbf{T} : \Omega \longrightarrow L(X)$ be holomorphic. Let $\sigma_\Omega(\mathbf{T})$ and $\rho_\Omega(\mathbf{T})$ be the spectrum and resolvent set of $\mathbf{T}(z)$, respectively. Suppose that $\Gamma \subset \rho_\Omega(\mathbf{T})$ and that $\sigma_0 := \sigma(\mathbf{T}) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\}$ consists of ℓ distinct discrete eigenvalues of $\mathbf{T}(z)$. We consider following moment problem for the eigenvalues.

Problem-C: *For $k \in \mathbb{Z}_+$, consider the tracial moment*

$$s_k := \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma z^k \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right).$$

Given the moments $s_0, \ldots, s_p$, for an appropriate $p \in \mathbb{N}$, determine the ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ and their algebraic multiplicities.

We show that exactly 2ℓ moments, namely, $s_0, \ldots, s_{2\ell-1}$ are needed in order to determine the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ and their algebraic multiplicities. We show that $\lambda_1, \ldots, \lambda_\ell$ are simple eigenvalues of an $\ell \times \ell$ Hankel matrix pencil $\widehat{H}_\ell - \lambda H_\ell$ with H_ℓ being nonsingular, where

$$H_\ell := \begin{bmatrix} s_0 & \cdots & s_{\ell-1} \\ \vdots & & \vdots \\ s_{\ell-1} & \cdots & s_{2\ell-2} \end{bmatrix} \text{ and } \widehat{H}_\ell := \begin{bmatrix} s_1 & \cdots & s_\ell \\ \vdots & & \vdots \\ s_\ell & \cdots & s_{2\ell-1} \end{bmatrix}.$$

Thus, when $\mathbf{T}(z)$ is an $n \times n$ holomorphic matrix, approximating the contour integral by numerical quadrature

$$s_p = \frac{1}{2\pi i}\int_\Gamma z^p \mathrm{Tr}\left(\mathbf{T}(z)^{-1}\mathbf{T}'(z)\right) dz \approx \sum_{j=1}^N w_j z_j^p \mathrm{Tr}\left(\mathbf{T}(z_j)^{-1}\mathbf{T}'(z_j)\right)$$

we obtain a numerical method for solving the eigenvalue problem $\mathbf{T}(\lambda)v = 0$.

For the recovery of eigenvectors, we proceed as follows. Set $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$. Let $P_j := \mathrm{Res}(\lambda_j, \mathbf{M})$ be the residue of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. We show that

$$R(P_j) = N(\mathbf{T}(\lambda_j))$$

when λ_j is a semisimple eigenvalue of $\mathbf{T}(z)$. Here $R(A)$ and $N(A)$ denote the range space and the null space of an operator A, respectively.

Problem-D: *For $k \in \mathbb{Z}_+$, consider the operator moment*

$$S_k := \frac{1}{2\pi i}\int_\Gamma z^k\, \mathbf{T}(z)^{-1} dz.$$

Given the moments $S_0, \ldots, S_p$, for an appropriate $p \in \mathbb{N}$, determine the residue P_j of $\mathbf{T}(z)^{-1}$ at λ_j for $j = 1 : \ell$.

We show that exactly ℓ operator moments, namely, $S_0, \ldots, S_{\ell-1}$, are needed in order to recover the residues $P_1, \ldots, P_\ell$ provided that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $\mathbf{T}(z)$. Indeed, let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$ and let v_j be given by $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then for $j = 1 : \ell$, we show that

$$\begin{aligned} \mathbf{P}_j &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)\mathbf{T}(z)^{-1}dz\right)/\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial and $\alpha_j := q_{v_j}(\lambda_j)$.

In particular, we show that if $v \in X$ and $\mathbf{v}_j := \begin{bmatrix} S_0 v \cdots S_{\ell-1} v \end{bmatrix} v_j$ then $\mathbf{v}_j$ is an eigenvector of $\mathbf{T}(z)$ corresponding to λ_j whenever $\mathbf{v}_j \neq 0$.

Obviously, when $\mathbf{T}(z)$ is a holomorphic matrix, approximating the contour integral in P_j by a numerical quadrature we obtain a numerical method for computing eigenvectors of $\mathbf{T}(z)$ from those of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$.

The thesis is organized as follows. We present preliminaries in the rest of Chapter-1. We undertake a detailed global spectral analysis of holomorphic and meromorphic matrices in Chapter-2. We present, among other things, global Smith forms of holomorphic matrices and global Smith-McMillan forms of meromorphic matrices. We also present system theoretic analysis of meromorphic matrices by considering global MFDs, transfer function realizations, and analytic system matrices and discuss their canonical forms. In Chapter-3, we develop spectral perturbation theory for holomorphic operator-valued functions. We discuss linearization of holomorphic operator-valued functions and present, among other things, perturbation bounds for discrete eigenvalues and their corresponding eigenvectors. Chapter-4 is devoted to local minimal realization of meromorphic matrices and linearization of meromorphic matrices. We show that contour integration based methods for solving holomorphic eigenvalue problems follow as special cases of local minimal realization of meromorphic matrices. We utilize local minimal realization to construct linearization of a meromorphic matrix which can be used for spectral perturbation analysis of meromorphic eigenvalue problems. In Chapter-5, we present spectral recovery problems for bounded linear operators and bounded operator pencils. We discuss recovery of eigenvalues from tracial moments and recovery of individual spectral projections from operator moments. Chapter-6 is devoted to spectral recovery problems for holomorphic matrix/operator-valued functions. We show that the spectral recovery methods naturally provide numerical methods for solving nonlinear eigenvalue problems. We illustrate performance of the numerical methods by considering a few numerical examples.

1.1 Preliminaries

Let $\mathbb{C}^n$ (resp., $\mathbb{R}^n$) denote the vector space of all column vectors $v := \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}^\top$ with $v_1, \ldots, v_n$ in $\mathbb{C}$ (resp., $\mathbb{R}$). We denote the vector space of all $m \times n$ matrices with entries in $\mathbb{C}$ (resp., $\mathbb{R}$) by $\mathbb{C}^{m\times n}$ (resp., $\mathbb{R}^{n\times n}$). Let $A \in \mathbb{C}^{m\times n}$. We denote the transpose of A by $A^\top$ and the conjugate transpose of A by A^*. The null space $N(A)$ and the range space $R(A)$ of A are given by

$$N(A) := \{v \in \mathbb{C}^n : Av = 0\} \text{ and } R(A) := \{Av : v \in \mathbb{C}^n\}.$$

We denote the identity matrix of size n by I_n and also by I when the size is clear from the context. We denote the j-th column of I_n by e_j for $j = 1 : n$.

Now, suppose that $A \in \mathbb{C}^{n\times n}$. Then $\lambda \in \mathbb{C}$ is said to be an eigenvalue of A if there exists a nonzero vector $v \in \mathbb{C}^n$ such that

$$Av = \lambda v.$$

The vector v is called an eigenvector of A corresponding to the eigenvalue λ. The eigenvector v is also called a right eigenvector. A nonzero vector $u \in \mathbb{C}^n$ is called a left eigenvector of A corresponding to λ if

$$u^\top A = \lambda u^\top.$$

The monic polynomial $p(z) := \det(zI_n - A)$ is the characteristic polynomial of A. We denote the spectrum of A by $\mathrm{eig}(A)$ which is given by

$$\mathrm{eig}(A) := \{\lambda \in \mathbb{C} : \mathrm{rank}(A - \lambda I_n) < n\} = \{\lambda \in \mathbb{C} : p(\lambda) = 0\}.$$

Let $\lambda \in \mathrm{eig}(A)$. The multiplicity λ as a root of $p(z)$ is called the algebraic multiplicity of λ and $g(\lambda) := \dim(A - \lambda I_n)$ is called the geometric multiplicity of λ. If $m(\lambda)$ is the algebraic multiplicity of λ then $m(\lambda) = \dim N((A - \lambda I_n)^n)$.

An eigenvalue λ of A is said to be simple if $m(\lambda) = 1$ and it is said to be semisimple if $m(\lambda) = g(\lambda)$. The multiplicity of λ as a root of the minimal polynomial of A is called the ascent of λ. Equivalently, $\nu(\lambda)$ is the ascent of λ as an eigenvalue of $A \Longleftrightarrow$ the order of the pole of $(A - zI_n)^{-1}$ at λ is $\nu(\lambda)$; see [39].

We write a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$ by $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Similarly, we write a block diagonal matrix with diagonal blocks $A_1, \ldots, A_m$ as

$$\mathrm{diag}(A_1, \ldots, A_m) = A_1 \oplus \cdots \oplus A_m.$$

An $m \times n$ block matrix (or a partitioned matrix) is a matrix of the form

$$A := \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

where each A_{ij} is a $p_i \times q_j$ matrix for $i = 1 : m$ and $j = 1 : n$.

Then $\begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$ is the i-th block row of A and $\begin{bmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{bmatrix}$ is the j-th block column of A.

Definition 1.1.1. *[52] Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$. The Kronecker product of A and B is defined as*

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \ldots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp\times nq}$$

where a_{ij} is (i, j)-th the entry of A.

The Kronecker product have the following properties:

- For $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{r\times s}, C \in \mathbb{C}^{n\times p}$, and $D \in \mathbb{C}^{s\times t}$,

$$(A \otimes B)(C \otimes D) = AC \otimes BD \in \mathbb{C}^{mr\times pt}.$$

- For any matrices A and B, $(A \otimes B)^\top = A^\top \otimes B^\top$ and $(A \otimes B)^* = A^* \otimes B^*$.
- If both A and B are non-singular, then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

Let A $\in \mathbb{C}^{mn\times mn}$ be an $m\times m$ block matrix with $n\times n$ blocks. Then $(e_j^\top \otimes I_n)$A is the j-th block row of A and A$(e_j\otimes I_n)$ is the j-th block column of A for $j=1:m$.

Singular Value Decomposition (SVD). A matrix $U\in\mathbb{C}^{n\times n}$ is said to be unitary if $U^*U = UU^* = I_n$. Let $A\in\mathbb{C}^{m\times n}$ and $\text{rank}(A)=r$. Then there exist unitary matrices $U\in\mathbb{C}^{m\times m}$ and $V\in\mathbb{C}^{n\times n}$ and real numbers $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0$ such that

$$A = U\left[\begin{array}{c|c} \text{diag}(\sigma_1,\ldots,\sigma_r) & 0 \\ \hline 0 & 0\end{array}\right]V^* = U\left[\begin{array}{c|c} \Sigma_r & 0 \\ \hline 0 & 0\end{array}\right]V^* = U\Sigma V^*. \tag{1.5}$$

The decomposition in (1.5) is called a singular value decomposition (SVD) of A. The positive numbers $\sigma_1,\ldots,\sigma_r$ are called nonzero singular values of A. The columns of V are called right singular vectors of A and columns of U are called left singular vectors of A. Let $U=\begin{bmatrix}u_1 & \cdots & u_m\end{bmatrix}$ and $V=\begin{bmatrix}v_1 & \cdots & v_n\end{bmatrix}$ be column partitions of A and B, respectively. Then $AV=U\Sigma$ yields

$$\begin{aligned} Av_j &= \sigma_j u_j,\ j=1:r &\implies& R(A)=\text{span}(u_1,\ldots,u_r)\\ Av_j &= 0,\ j=r+1:n &\implies& N(A)=\text{span}(v_{r+1},\ldots,v_n)\end{aligned}$$

and $A^*U = V\Sigma^*$ yields

$$\begin{aligned} A^*u_j &= \sigma_j v_j,\ j=1:r &\implies& R(A^*)=\text{span}(v_1,\ldots,v_r)\\ A^*u_j &= 0,\ j=r+1:m &\implies& N(A^*)=\text{span}(u_{r+1},\ldots,u_m).\end{aligned}$$

Set $U_r := \begin{bmatrix}u_1 & \cdots & u_r\end{bmatrix}$ and $V_r := \begin{bmatrix}V_1 & \cdots & v_r\end{bmatrix}$. Then $U_r\in\mathbb{C}^{m\times r}$ and $V_r\in\mathbb{C}^{n\times r}$ are isometry, that is, $U_r^*U_r = I_r$ and $V_r^*V_r = I_r$. Then the SVD in (1.5) can be rewritten as

$$A = U_r\Sigma_r V_r^* = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^*. \tag{1.6}$$

The SVD in (1.6) is called a compact SVD or a trimmed SVD of A.

Definition 1.1.2. *[30] Let $s_0, s_1, \cdots s_{2n}$ be a sequence. An $n\times n$ matrix of the form*

$$H_n = \begin{bmatrix} s_0 & s_1 & \cdots & s_{n-2} & s_{n-1}\\ s_1 & s_2 & \cdots & s_{n-1} & s_n\\ \vdots & \vdots & & \vdots & \vdots\\ s_{n-2} & s_{n-1} & \cdots & s_{2n-4} & s_{2n-3}\\ s_{n-1} & s_n & \cdots & s_{2n-3} & s_{2n-2}\end{bmatrix} = \text{Hankel}(s_0,\ldots,s_{2n-2})$$

is called a Hankel matrix. The matrix

$$\widehat{H}_n = \begin{bmatrix} s_1 & s_2 & \cdots & s_{n-1} & s_n \\ s_2 & s_3 & \cdots & s_n & s_{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-1} & s_{2n-2} \\ s_n & s_{n+1} & \cdots & s_{2n-2} & s_{2n-1} \end{bmatrix} = \text{Hankel}(s_1, \ldots, s_{2n-1})$$

is called a shifted Hankel matrix.

Observe that the (i, j)-th entry of H_n is given by $a_{ij} = s_{i+j-2}$ and the (i, j)-th entry of $\widehat{H}_n$ is given by $a_{ij} = s_{i+j-1}$. The entries of a Hankel matrix are constant along the diagonals perpendicular to the main diagonal.

Polynomial/rational vectors and matrices. Let $\mathbb{C}[z]$ denote the ring of scalar polynomials with coefficients in $\mathbb{C}$ and $\mathbb{C}(z)$ denote the field of rational functions of the form $p(z)/q(z)$, where $p(z)$ and $q(z)$ are polynomials in $\mathbb{C}[z]$. We denote by $\mathbb{C}[z]^{m\times n}$ and $\mathbb{C}(z)^{m\times n}$ the sets of all $m \times n$ matrices with entries in $\mathbb{C}[z]$ and $\mathbb{C}(z)$, respectively. The elements of $\mathbb{C}[z]^{m\times n}$ are called matrix polynomials. If $P(z) \in \mathbb{C}[z]^{m\times n}$ then it can be written as

$$P(z) = A_0 + zA_1 + \cdots + z^\ell A_\ell$$

for some $m \times n$ matrices $A_0, \ldots, A_\ell$ such that $A_\ell \neq 0$. The number ℓ is called the degree of $P(z)$. We write $\ell = \deg(P(z))$. The elements of $\mathbb{C}(z)^{m\times n}$ are called rational matrices.

We denote $\mathbb{C}[z]^{n\times 1}$ by $\mathbb{C}[z]^n$ and $\mathbb{C}(z)^{n\times 1}$ by $\mathbb{C}(z)^n$. Thus $\mathbb{C}[z]^n$ is a free module over the ring $\mathbb{C}[z]$ and a vector space of $\mathbb{C}$. On the other hand, $\mathbb{C}(z)^n$ is a vector space over the field $\mathbb{C}(z)$ and is called a rational vector space. We refer to elements of $\mathbb{C}[z]^n$ as polynomial vectors (also as vector polynomials) and refer to elements of $\mathbb{C}(z)^n$ as rational vectors.

Let $P(z) \in \mathbb{C}[z]^{m\times n}$. Then $P(z) : \mathbb{C}[z]^n \longrightarrow \mathbb{C}[z]^m,\ v(z) \longmapsto P(z)v(z)$, is a module homomorphism. The rank of $P(z)$ as a module homomorphism is called the ***normal rank*** of $P(z)$ and is denoted by $\text{nrank}(P)$. It can be shown that

$$\text{nrank}(P) = \max\{\text{rank}(P(\lambda)) : \lambda \in \mathbb{C}\}.$$

The matrix polynomial $P(z)$ is said to be ***regular*** if $m = n$ and $\text{nrank}(P) = n$. If $P(z)$ is not regular then $P(z)$ is said to be ***singular.***

Definition 1.1.3 (eigenvalue). *Let $P(z) \in \mathbb{C}[z]^{m\times n}$. A complex number $\lambda \in \mathbb{C}$ is said to be an* eigenvalue *of $P(z)$ if* $\mathrm{rank}(P(\lambda)) < \mathrm{nrank}(P)$. *The spectrum of $P(z)$ is given by* $\sigma_{\mathbb{C}}(P) := \{\lambda \in \mathbb{C} : \mathrm{rank}(P(\lambda)) < \mathrm{nrank}(P)\}$.

A matrix polynomial $U(z) \in \mathbb{C}[z]^{n\times n}$ is said to be ***unimodular*** if $\det(U(z))$ is a nonzero constant for all $z \in \mathbb{C}$. Thus, if $U(z)$ is unimodular then $U(z)^{-1}$ is again a matrix polynomial. Let $A(z), B(z) \in \mathbb{C}[z]^{m\times n}$. Then $A(z)$ and $B(z)$ are said to be ***unimodularly*** equivalent if there exist unimodular matrices $U(z) \in \mathbb{C}[z]^{m\times m}$ and $V(z) \in \mathbb{C}[z]^{n\times n}$ such that $U(z)A(z)V(z) = B(z)$.

The Smith canonical form [25, 32, 56, 50] of $P(z)$ provides a complete information about the eigen-structure of $P(z)$.

Theorem 1.1.4 (Smith form, [25]). *Let $P(z) \in \mathbb{C}[z]^{m\times n}$ and $r = \mathrm{nrank}(P)$. Then there exist unimodular matrices $U(z) \in \mathbb{C}[z]^{m\times m}$ and $V(z) \in \mathbb{C}[z]^{n\times n}$ such that*

$$U(z)P(z)V(z) = \left[\begin{array}{c|c} \mathrm{diag}(\phi_1(z), \ldots, \phi_r(z)) & 0 \\ \hline 0 & 0 \end{array}\right] =: S_P(z),$$

where $\phi_1(z), \ldots, \phi_r(z)$ are monic polynomials in $\mathbb{C}[z]$ such that $\phi_i(z)$ divides ϕ_{i+1} for $i = 1 : r-1$.

The diagonal matrix $S_P(z)$ is called the ***Smith form*** or Smith canonical form of $P(z)$. The monic polynomials $\phi_1(z), \ldots, \phi_r(z)$ are called invariant polynomials of $P(z)$. The polynomial $\phi_P(z) := \prod_{j=1}^{r} \phi_j(z)$ is called the zero polynomial of $P(z)$. It is easy to see that λ is an eigenvalue of $P(z) \iff \phi_P(\lambda) = 0$. Thus $\sigma_{\mathbb{C}}(P) = \{\lambda \in \mathbb{C} : \phi_P(\lambda) = 0\}$.

Let λ be an eigenvalue of $P(z)$. Then $\phi_P(\lambda) = 0$. The multiplicity of λ as a root of $\phi_P(z)$ is called the ***algebraic multiplicity*** of λ as an eigenvalue of $P(z)$. For $j = 1 : r$, we have $\phi_i(z) = (z-\lambda)^{m_i}\psi_i(z)$ with $\psi_i(\lambda) \neq 0$ and $m_i \geq 0$. Since $\phi_i(z)$ divides ϕ_{i+1} for $i = 1 : r-1$, it follows that $0 \leq m_1 \leq \cdots \leq m_r$. The tuple

$$\mathrm{Ind}(\lambda, P) := (m_1, \ldots, m_r)$$

is called the ***index*** of the eigenvalue λ. Then $m := m_1 + \cdots + m_r$ is the algebraic multiplicity of the eigenvalue λ of $P(z)$. The factors $(z-\lambda)^{m_i}$ with $m_i \geq 1$ are called the elementary divisors of $P(z)$ at λ. Suppose that

$$0 = m_1 = \cdots = m_{\ell-1} < m_\ell \leq \cdots \leq m_r.$$

Then $m_\ell, \ldots, m_r$ are called the ***partial multiplicities*** of the eigenvalue λ of $P(z)$ and $\phi_\ell(z), \ldots, \phi_r(z)$ are called the ***non-unit invariant polynomials*** of $P(z)$.

Linearization is a classical technique widely used for reducing a polynomial eigenvalue problem to a generalized eigenvalue problem. It reduces a matrix polynomial $P(z) \in \mathbb{C}[z]^{n\times n}$ of degree m to an $mn \times mn$ matrix pencil of the form $L(z) = A + zB$ while preserving the eigenvalues and their partial multiplicities.

Definition 1.1.5 (Linearization, [25, 41])**.** *Let $P(z)$ be an $n \times n$ matrix polynomial of degree m. A pencil $L(z) := A + zB$ with $A, B \in \mathbb{C}^{mn\times mn}$ is called a linearization of $P(z)$ if there exist $mn \times mn$ unimodular matrix polynomials $U(z)$ and $V(z)$ such that*

$$U(z)L(z)V(z) = \left[\begin{array}{c|c} I_{(m-1)n} & 0 \\ \hline 0 & P(z) \end{array}\right].$$

It follows that $L(z)$ and $P(z)$ have the same non-unit invariant polynomials. Hence we have $\sigma_{\mathbb{C}}(P) = \sigma_{\mathbb{C}}(L)$. In particular, if $P(z)$ is regular then obviously $L(z)$ is regular and $\det(P(z)) = c\det(L(z))$ for some nonzero constant c independent of z.

Let $G(z) \in \mathbb{C}(z)^{m\times n}$. Then $G(z) : \mathbb{C}(z)^n \longrightarrow \mathbb{C}(z)^m,\ v(z) \longmapsto G(z)v(z)$, is a linear transformation over the field $\mathbb{C}(z)$. The rank of $G(z)$ as a linear transformation over the field $\mathbb{C}(z)$ is called the ***normal rank*** of $G(z)$ and is denoted by $\mathrm{nrank}(G)$. Define

$$\wp_{\mathbb{C}}(G) := \{\lambda \in \mathbb{C} : \lambda \text{ is a pole of } G(z)\}.$$

It can be shown that $\mathrm{nrank}(G) = \max\{\mathrm{rank}(G(\lambda)) : \lambda \in \mathbb{C} \text{ and } \lambda \notin \wp_{\mathbb{C}}(G)\}$. The rational matrix $G(z)$ is said to be ***regular*** if $m = n$ and $\mathrm{nrank}(G) = n$. Otherwise, $G(z)$ is said to be ***singular.*** The null space $N(G)$ and the range space $R(G)$ are given by

$$N(G) := \{v(z) \in \mathbb{C}(z)^n : G(z)v(z) = 0\} \text{ and } R(G) := \{G(z)v(z) : v(z) \in \mathbb{C}(z)^n\}.$$

Definition 1.1.6 (eigenvalue, zero and eigenpole, [2])**.** *Let $G(z) \in \mathbb{C}(z)^{m\times n}$.*

(a) A complex number $\lambda \in \mathbb{C}$ is said to be an eigenvalue of $G(z)$ if $\lambda \notin \wp_{\mathbb{C}}(G)$ and $\mathrm{rank}(G(\lambda)) < \mathrm{nrank}(G)$. The eigenspectrum of $G(z)$ is given by

$$\mathrm{eig}_{\mathbb{C}}(G) := \{\lambda \in \mathbb{C} : \lambda \notin \wp_{\mathbb{C}}(G) \text{ and } \mathrm{rank}(G(\lambda)) < \mathrm{nrank}(G)\}.$$

(b) A complex number λ is said to be a zero of $G(z)$ if there exists $v(z) \in \mathbb{C}[z]^n$ with $v(\lambda) \neq 0$ such that $\lim_{z\to\lambda} G(z)v(z) = 0$ and $v \notin N(G)$. The spectrum of $G(z)$ is given by $\sigma_{\mathbb{C}}(G) := \{\lambda \in \mathbb{C} : \lambda$ is an zero of $G(z)\}$.

(c) *A complex number* λ *is said to be an* eigenpole *of* $G(z)$ *if* λ *is a pole of* $G(z)$ *and there exists* $v(z) \in \mathbb{C}[z]^n$ *with* $v(\lambda) \neq 0$ *such that* $\lim_{z\to\lambda} G(z)v(z) = 0$ *and* $v \notin N(G)$. *Define* $\mathrm{eip}_{\mathbb{C}}(G) := \{\lambda \in \wp_{\mathbb{C}}(G) : \lambda$ *is an eigenpole of* $G(z)\}$.

The Smith-McMillan form [32, 56, 50] of $G(z)$ provides information about the zero and pole structures of $G(z)$.

Theorem 1.1.7 (Smith-McMillan form, [32, 50]). *Let* $G(z) \in \mathbb{C}(z)^{m\times n}$ *and* $r = \mathrm{nrank}(G)$. *Then there exist unimodular matrices* $U(z) \in \mathbb{C}[z]^{m\times m}$ *and* $V(z) \in \mathbb{C}[z]^{n\times n}$ *such that*

$$U(z)G(z)V(z) = \left[\begin{array}{c|c} \mathrm{diag}(\phi_1(z)/\psi_1(z), \ldots, \phi_r(z)/\psi_r(z)) & 0 \\ \hline 0 & 0 \end{array}\right] =: \Sigma_G(z),$$

where $\phi_1(z), \ldots, \phi_r(z)$ *and* $\psi_1(z), \ldots, \psi_r(z)$ *are monic polynomials in* $\mathbb{C}[z]$ *and* $\phi_i(z)$ *and* $\psi_i(z)$ *are coprime for* $i = 1 : r$. *Further,* $\phi_i(z)$ *divides* $\phi_{i+1}(z)$ *and* $\psi_{i+1}(z)$ *divides* $\psi_i(z)$ *for* $i = 1 : r-1$.

The diagonal matrix $\Sigma_G(z)$ is called the ***Smith-McMillan*** form of $G(z)$. The polynomials $\phi_1(z), \ldots, \phi_r(z)$ are called the ***invariant zero polynomials*** of $G(z)$ and $\psi_1(z), \ldots, \psi_r(z)$ are called the ***invariant pole polynomials*** of $G(z)$. The ***zero polynomial*** $\phi_G(z)$ and the ***pole polynomial*** $\psi_G(z)$ of $G(z)$ are given by

$$\phi_G(z) := \prod_{j=1}^{r} \phi_j(z) \text{ and } \psi_G(z) := \prod_{j=1}^{r} \psi_j(z).$$

It can be shown [2] that $\sigma_{\mathbb{C}}(G) = \mathrm{eig}_{\mathbb{C}}(G) \cup \mathrm{eip}_{\mathbb{C}}(G)$. In fact,

$$\sigma_{\mathbb{C}}(G) = \{\lambda \in \mathbb{C} : \phi_G(\lambda) = 0\}, \quad \wp_{\mathbb{C}}(G) = \{\lambda \in \mathbb{C} : \psi_G(\lambda) = 0\}$$

and $\mathrm{eip}_{\mathbb{C}}(G) = \{\lambda \in \mathbb{C} : \phi_G(\lambda) = 0 \text{ and } \psi_G(\lambda) = 0\}$. Index and partial multiplicities of a zero (resp., pole) of $G(z)$ can be defined by factorizing the invariant zero (resp., pole) polynomials $\phi_1(z), \ldots, \phi_r(z)$ (resp., $\psi_1(z), \ldots, \psi_r(z)$).

If $f(z) \in \mathbb{C}(z)$ then there exist $p(z), q(z) \in \mathbb{C}(z)$ such that $p(z)$ and $q(z)$ are coprime and $f(z) = p(z)/q(z)$. A similar decomposition of a rational matrix $G(z) \in \mathbb{C}(z)^{m\times n}$ is called matrix-fraction description (MFD).

Definition 1.1.8 (MFD, [32]). *A right matrix-fraction description (MFD) of a rational matrix* $G(z) \in \mathbb{C}(z)^{m\times n}$ *a decomposition of* $G(z)$ *of the form* $G(z) = N(z)D(z)^{-1}$, *where* $N(z) \in \mathbb{C}[z]^{m\times n}$ *and* $D(z) \in \mathbb{C}[z]^{n\times n}$ *is regular. A right MFD* $G(z) = N(z)D(z)^{-1}$ *is said to be right coprime if the common right divisors of* $N(z)$ *and* $D(z)$ *are unimodular matrix polynomials.*

Every rational matrix admits a right coprime MFD. Further, if $G(z) = N(z)D(z)^{-1}$ is a right coprime MFD then $\sigma_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(N)$ and $\wp_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(D)$, see [32, 56].

A rational matrix $G(z) \in \mathbb{C}(z)^{m\times n}$ is said to be ***proper*** if $G(z) \to D$ as $z \to \infty$ for some $D \in \mathbb{C}^{m\times n}$. If $G(z) \to 0$ as $z \to \infty$ then $G(z)$ is said to be ***strictly proper***. If $G(z)$ is not proper then it is said to be ***nonproper***.

Let $G(z) \in \mathbb{C}(z)^{m\times n}$ be strictly proper. Then $(C, A, B) \in \mathbb{C}^{m\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$ is said to be a ***realization*** of $G(z)$ if $G(z) = C(zI_r - A)^{-1}B$. The realization (C, A, B) is said to be ***minimal*** if

$$\operatorname{rank}\left(\begin{bmatrix} zI_r - A \\ C \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} zI_r - A & B \end{bmatrix}\right) = r$$

for all $z \in \mathbb{C}$. If $G(z)$ is nonproper then there exist a unique matrix polynomial $P(z)$ and a unique strictly proper rational matrix $G_{sp}(z)$ such that

$$G(z) = P(z) + G_{sp}(z).$$

If (C, A, B) is a realization of $G_{sp}(z)$ then $G(z) = P(z) + C(zI_r - A)^{-1}B$ is called a realization of $G(z)$. The matrix polynomial

$$S(z) := \left[\begin{array}{c|c} P(z) & C \\ \hline B & A - zI_r \end{array}\right]$$

is called the ***Rosenbrock system matrix*** and $G(z) = P(z) + C(zI_r - A)^{-1}B$ is called the ***transfer function*** of $S(z)$. The system matrix $S(z)$ is called ***irreducible*** if the realization (C, A, B) is minimal. If the system matrix $S(z)$ is irreducible then $\sigma_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(S)$ and $\wp_{\mathbb{C}}(G) = \operatorname{eig}(A)$. We refer to [32, 50, 56] for more on realization and Rosenbrock system matrix.

2

Spectral Analysis of Holomorphic and Meromorphic Matrix-Valued Functions

Consider once again the delay-differential systems (DDS) [18, 19, 59]

$$\begin{aligned}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} &= A_0\mathbf{x}(t) + \sum_{j=1}^{N_1} A_j\mathbf{x}(t-\tau_j) + \sum_{j=1}^{N_2} B_j\mathbf{u}(t-t_j),\\ \mathbf{y}(t) &= \sum_{j=1}^{N_3} C_j\mathbf{x}(t-s_j) + \sum_{j=1}^{N_4} D_j\mathbf{u}(t-h_j),\end{aligned}$$

where $\mathbf{x}(t) \in \mathbb{C}^r$ is a vector of state variables and $\mathbf{u}(t) \in \mathbb{C}^n$ is a vector of control variables. For $h \in [0, \infty)$, consider the shift operators

$$S_h : \mathbf{x}(t) \longmapsto \mathbf{x}(t-h) \text{ and } T_h : \mathbf{u}(t) \longmapsto \mathbf{u}(t-h).$$

Then for a nonzero $\mathbf{x} \in \mathbb{C}^r$ and $\lambda \in \mathbb{C}$, it follows that $\mathbf{x}e^{\lambda t}$ is an eigen-function of S_h corresponding to the eigenvalue $e^{-\lambda h}$. Indeed, we have $S_h\mathbf{x}e^{\lambda t} = e^{-\lambda h}\mathbf{x}e^{\lambda t}$. Similarly, for a nonzero $\mathbf{u} \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, we have $T_h\mathbf{u}e^{\lambda t} = e^{-\lambda h}\mathbf{u}e^{\lambda t}$. Hence, for the ansatz $\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{u}(t)\end{bmatrix} := \begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} e^{\lambda t}$ and $\begin{bmatrix}\mathbf{x}(t)\\ \mathbf{u}(t)\end{bmatrix} := \begin{bmatrix}\mathbf{x}\\ -\mathbf{u}\end{bmatrix} e^{\lambda t}$, respectively, the DDS yields

$$\begin{bmatrix}-A(\lambda) & B(\lambda)\\ C(\lambda) & D(\lambda)\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} e^{\lambda t} = \begin{bmatrix}0\\ \mathbf{y}(t)\end{bmatrix} \text{ and } \begin{bmatrix}A(\lambda) & B(\lambda)\\ -C(\lambda) & D(\lambda)\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ -\mathbf{u}\end{bmatrix} e^{\lambda t} = \begin{bmatrix}0\\ -\mathbf{y}(t)\end{bmatrix} \tag{2.1}$$

where $A(\lambda) := \lambda I_r - A_0 - \sum_{j=1}^{N_1} A_j e^{-\lambda\tau_j}$, $B(\lambda) := \sum_{j=1}^{N_2} B_j e^{-\lambda t_j}$, $C(\lambda) := \sum_{j=1}^{N_3} C_j e^{-\lambda s_j}$ and $D(\lambda) := \sum_{j=1}^{N_4} D_j e^{-\lambda h_j}$. The holomorphic matrices

$$\mathbf{S}(\lambda) := \begin{bmatrix}-A(\lambda) & B(\lambda)\\ C(\lambda) & D(\lambda)\end{bmatrix} \text{ and } \mathbf{H}(\lambda) := \begin{bmatrix}A(\lambda) & B(\lambda)\\ -C(\lambda) & D(\lambda)\end{bmatrix}$$

are called system matrices of the DDS. This shows that multiple system matrices can be associated with the same DDS which, as we shall see, are equivalent in an appropriate sense. The meromorphic matrix

$$\mathbf{M}(\lambda) := D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$$

is called the transfer function of the DDS. Eliminating $\mathbf{x}e^{\lambda t}$ in (2.1, we have

$$\mathbf{y}(t) = [D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)]\mathbf{u}e^{\lambda t} = \mathbf{M}(\lambda)\mathbf{u}e^{\lambda t}.$$

Observe that $\mathbf{y}(t) = 0$ whenever $\mathbf{M}(\lambda)\mathbf{u} = 0$, that is, when $(\lambda, \mathbf{u})$ is an eigenpair of $\mathbf{M}(z)$. Alternatively, it follows from (2.1) that $\mathbf{y}(t) = 0$ whenever

$$\mathbf{S}(\lambda)\begin{bmatrix}\mathbf{x}\\ \mathbf{u}\end{bmatrix} = 0 \text{ or } \mathbf{H}(\lambda)\begin{bmatrix}\mathbf{x}\\ -\mathbf{u}\end{bmatrix} = 0.$$

This shows that the zero output of the DDS can be analyzed via eigenvalues and eigenvectors of the system matrix as well as eigenvalues and eigenvectors of the transfer function. This motivates us to study the spectral analysis of holomorphic and meromorphic matrix-valued functions.

The Kronecker canonical form (KCF) of an $m \times n$ matrix pencil $A - \lambda B$ provides a complete information about the spectral structure of the pencil. The KCF states that [25, 32] there exist nonsingular matrices P and Q such that

$$P(A - \lambda B)Q = \operatorname{diag}(L_{\varepsilon_1}, L_{\varepsilon_2}, \ldots, L_{\varepsilon_p}, L_{\eta_1}^T, L_{\eta_2}^T, \ldots, L_{\eta_s}^T, \lambda N - I_p, J - \lambda I_\ell),$$

where N is nilpotent, both N and J are in Jordan canonical form and L_k is the $k \times (k+1)$ bidiagonal pencil given by

$$L_k = \begin{bmatrix} -\lambda & 1 & & & \\ & -\lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & -\lambda & 1 \end{bmatrix}.$$

The elementary divisors of $J - \lambda I$ and $\lambda N - I$, respectively, are the finite and infinite elementary divisors of the pencil $A - \lambda B$ and constitute the regular part of $A - \lambda B$. The L_k and $L_k^\top$ blocks constitute the singular part of $A - \lambda B$. The index sets $\{\varepsilon_1, \ldots, \varepsilon_p\}$ and $\{\eta_1, \ldots, \eta_s\}$ are the right and left minimal indices (Kronecker indices) of $A - \lambda B$, respectively.

A KCF like canonical form of an $m \times n$ matrix polynomial $P(z)$ with $\deg(P(z)) \geq 2$ is not available. Nevertheless, the Smith canonical form [25, 32, 50, 56] of $P(z)$ provides a complete spectral information. On the other hand, the Smith-McMillan form [32, 50, 56, 9] provides a complete spectral information of an $m \times n$ rational matrix.

Let $\Omega \subset \mathbb{C}$ be a connected open set. For a holomorphic matrix-valued function $\mathbf{H} : \Omega \longrightarrow \mathbb{C}^{m\times n}$, a local version of the Smith canonical form has been considered in [24, 9, 35, 44] when Ω is a compact region and $\mathbf{H}(z)$ is regular. Also, for a meromorphic matrix-valued function $\mathbf{M} : \Omega \longrightarrow \mathbb{C}^{m\times n}$, a local version of the Smith-McMillan form has been considered in [9, 26, 35] when Ω is a compact region and $\mathbf{M}(z)$ is regular. On the other hand, the existence of a global canonical form of $\mathbf{M}(z)$ has been proved in [38] where it is shown that $\mathbf{M}(z)$ is *equivalent* to a diagonal matrix $\mathrm{diag}(f_1(z), \ldots, f_r(z), 0, \ldots, 0)$, where $f_1(z), \ldots, f_r(z)$ are meromorphic functions on Ω such that $f_i(z)$ divides $f_{i+1}(z)$ for $i = 1 : r-1$. See also [29].

Although, local canonical forms such as local Smith forms of holomorphic and meromorphic matrices have been studied extensively in the literature, see [24, 9, 26, 35, 44, 38, 29] and the references therein, there has been relatively little research on global canonical forms of holomorphic and meromorphic matrices. One of the main aims of this chapter is to present a comprehensive treatment of global canonical forms of holomorphic and meromorphic matrix-valued functions along with their detailed proofs. We present global Smith forms and Smith-McMillan forms of holomorphic and meromorphic matrix-valued functions which are akin to Smith forms and Smith-McMillan forms of matrix polynomials and rational matrices, respectively. We utilize Weierstrass products and Weierstrass factorization of holomorphic functions to prove global Smith forms and Smith-McMillan forms of holomorphic and meromorphic matrix-valued functions.

The second part of the chapter is devoted to system theoretic analysis of meromorphic matrix-valued functions. We analyze matrix-fraction description (MFD) and transfer function realization of a meromorphic matrix-valued function. We prove, among other things, relation between canonical forms of system matrices and their transfer functions.

2.1 Canonical forms of holomorphic matrix-valued functions

Let $\mathcal{O} \subset \mathbb{C}$ be a **domain**, that is, $\mathcal{O}$ is a nonempty connected open set. Let X be a complex Banach space. A function $f : \mathcal{O} \longrightarrow X$ is said to be holomorphic (or analytic) in $\mathcal{O}$ if f is differentiable on $\mathcal{O}$. We denote the set of all X-valued holomorphic functions on $\mathcal{O}$ by $\mathbb{H}(\mathcal{O}, X)$, that is,

$$\mathbb{H}(\mathcal{O}, X) := \{f : \mathcal{O} \longrightarrow X \mid f \text{ is holomorphic in } \mathcal{O}\}.$$

Let $\Omega \subset \mathbb{C}$ be a **region**, that is, the interior of Ω is a domain and Ω possibly contains a part or all of its boundary $\partial\Omega$. A function $f : \Omega \longrightarrow X$ is said to be holomorphic in Ω if $f \in \mathbb{H}(\mathcal{O})$ for some open set $\mathcal{O} \subset \mathbb{C}$ such that $\Omega \subset \mathcal{O}$. We denote the set of all X-valued holomorphic functions on Ω by $\mathbb{H}(\Omega, X)$.

For simplicity of notation, when $X = \mathbb{C}$ and $X = \mathbb{C}^{m\times n}$, we set

$$\mathbb{H}(\Omega) := \mathbb{H}(\Omega, \mathbb{C}),\ \mathbb{H}(\Omega)^n := \mathbb{H}(\Omega, \mathbb{C}^n) \text{ and } \mathbb{H}(\Omega)^{m\times n} := \mathbb{H}(\Omega, \mathbb{C}^{m\times n}).$$

We refer to the elements of $\mathbb{H}(\Omega)^n$ as **holomorphic vectors** and the elements of $\mathbb{H}(\Omega)^{m\times n}$ as **holomorphic matrices**. Note that $\mathbb{H}(\Omega)$ is a commutative ring of holomorphic functions on Ω and $\mathbb{H}(\Omega)^{n\times n}$ is a non-commutative ring of $n \times n$ holomorphic matrix-valued functions on Ω.

Definition 2.1.1. *Let $A \in \mathbb{H}(\Omega)^{n\times n}$. Then A is said to be* **invertible** *(or a unit element of $\mathbb{H}(\Omega)^{n\times n}$) if there exists $B \in \mathbb{H}(\Omega)^{n\times n}$ such that $AB = BA = I$, that is,*

$$A(z)B(z) = B(z)A(z) = I_n \quad \text{for all} \quad z \in \Omega.$$

In such a case, B is called an inverse of A.

Remark 2.1.2. *If $A \in \mathbb{H}(\Omega)^{n\times n}$ is invertible then A has a unique inverse which is denoted by A^{-1}. If $A(z)$ is invertible for all $z \in \Omega$ then defining $A^{-1}(z) := A(z)^{-1}$ for $z \in \Omega$, it follows that $A^{-1} \in \mathbb{H}(\Omega)^{n\times n}$ and $AA^{-1} = A^{-1}A = I$, that is, A^{-1} is the unique inverse of A.*

Let $\mathrm{GL}_n(\mathbb{H}(\Omega))$ denote the group of invertible elements in $\mathbb{H}(\Omega)^{n\times n}$, that is,

$$\mathrm{GL}_n(\mathbb{H}(\Omega)) := \{A \in \mathbb{H}(\Omega)^{n\times n} \mid A \text{ is invertible }\}.$$

The elements in $\mathrm{GL}_n(\mathbb{H}(\Omega)) \subset \mathbb{H}(\Omega)^{n\times n}$ are called units or unit elements of $\mathbb{H}(\Omega)^{n\times n}$.

Definition 2.1.3. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then A is said to be* **regular** *if $m = n$ and the matrix $A(z)$ is invertible for some $z \in \Omega$. If A is not regular then A is said to be* **singular.**

For a regular matrix-valued function, the spectrum is defined as follows.

Definition 2.1.4. *Let $A \in \mathbb{H}(\Omega)^{n\times n}$ be regular. The spectrum $\sigma_\Omega(A)$ and the resolvent set $\rho_\Omega(A)$ of A are given by*

$$\sigma_\Omega(A) := \{z \in \Omega : \det(A(z)) = 0\} \text{ and } \rho_\Omega(A) = \Omega \setminus \sigma_\Omega(A),$$

where $\det(M)$ is the determinant of a matrix $M \in \mathbb{C}^{n\times n}$.

Note that $A \in \mathrm{GL}_n(\mathbb{H}(\Omega)) \iff \sigma_\Omega(A) = \emptyset$. For a nonzero scalar function $f \in \mathbb{H}(\Omega)$ the spectrum $\sigma_\Omega(f)$ is the set of zeros of f, that is, $\sigma_\Omega(f) = \{z \in \Omega : f(z) = 0\}$. Thus $\sigma_\Omega(f) = \emptyset \iff f$ is a unit element of $\mathbb{H}(\Omega) \iff 1/f \in \mathbb{H}(\Omega)$. It is well known [58] that $\sigma_\Omega(f)$ is at most a countable set with no accumulation points in Ω. Consequently, $\sigma_\Omega(A)$ is at most a countable set with no accumulation points in Ω when $A \in \mathbb{H}(\Omega)^{n\times n}$ is regular.

For spectral analysis, we view a holomorphic matrix as a homomorphism between modules. Therefore, we briefly discuss some basic facts about $\mathbb{H}(\Omega)$-modules. Note that $\mathbb{H}(\Omega)^n$ is a module over $\mathbb{H}(\Omega)$. The span of a subset $\{f_1, \ldots, f_\ell\} \subset \mathbb{H}(\Omega)^n$ is given by

$$\mathrm{span}_{\mathbb{H}(\Omega)}(f_1, \ldots, f_\ell) := \{g_1 f_1 + \cdots + g_\ell f_\ell : \;\; g_1, \ldots, g_\ell \in \mathbb{H}(\Omega)\}.$$

Obviously, $\mathrm{span}_{\mathbb{H}(\Omega)}(f_1, \ldots, f_\ell)$ is a submodule of $\mathbb{H}(\Omega)^n$. A subset $\{f_1, \ldots, f_\ell\} \subset \mathbb{H}(\Omega)^n$ is said to be **linearly dependent** if there exist $g_1, \ldots, g_\ell$ in $\mathbb{H}(\Omega)$ not all are zero such that $g_1 f_1 + \cdots + g_\ell f_\ell = 0$. Thus $\{f_1, \ldots, f_\ell\}$ is said to be **linearly independent** if it is not linearly dependent, that is, if $g_1 f_1 + \cdots + g_\ell f_\ell = 0$ then $g_1 = \cdots = g_\ell = 0$.

Definition 2.1.5. *Let $\mathcal{V} \subset \mathbb{H}(\Omega)^n$ be a submodule. Then $\mathcal{B} \subset \mathbb{H}(\Omega)^n$ is said to be a* **basis** *of $\mathcal{V}$ if $\mathcal{B}$ is linearly independent and $\mathrm{span}_{\mathbb{H}(\Omega)}(\mathcal{B}) = \mathcal{V}$. The dimension of $\mathcal{V}$ is the number of elements in $\mathcal{B}$ and is denoted by $\dim_{\mathbb{H}(\Omega)}(\mathcal{V})$.*

Obviously $\{e_1, \ldots, e_n\}$ is a basis of the module $\mathbb{H}(\Omega)^n$ and $\dim_{\mathbb{H}(\Omega)}(\mathbb{H}(\Omega)^n) = n$. In fact, $\mathbb{H}(\Omega)^n$ is a free module. It is easily seen that if $\mathcal{V} \subset \mathbb{H}(\Omega)^n$ is a submodule and $\mathcal{V} \neq \{0\}$ then $\mathcal{V}$ has a basis and $\dim_{\mathbb{H}(\Omega)}(\mathcal{V}) \leq n$.

Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then for each fixed $\omega \in \Omega$, we can talk about the rank of the matrix $A(\omega) \in \mathbb{C}^{m\times n}$ over the field $\mathbb{C}$, which we denote by $\mathrm{rank}(A(\omega))$ or $\mathrm{rank}_\mathbb{C}(A(\omega))$. Now consider A as a module homomorphism

$$A : \mathbb{H}(\Omega)^n \longrightarrow \mathbb{H}(\Omega)^m, f \longmapsto Af.$$

The rank of A as a module homomorphism is called the ***normal rank*** of A. Indeed, consider

$$N(A) := \{f \in \mathbb{H}(\Omega)^n : Af = 0\} \subset \mathbb{H}(\Omega)^n \text{ and } R(A) := \{Af : f \in \mathbb{H}(\Omega)^n\} \subset \mathbb{H}(\Omega)^m.$$

Then $N(A)$ is a submodule of $\mathbb{H}(\Omega)^n$ and $R(A)$ is a submodule of $\mathbb{H}(\Omega)^m$.

Definition 2.1.6 (normal rank). *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then the rank of A is given by*

$$\operatorname{rank}_{\mathbb{H}(\Omega)}(A) := \dim_{\mathbb{H}(\Omega)}(R(A)).$$

The rank $\operatorname{rank}_{\mathbb{H}(\Omega)}(A)$ is also called the normal rank of A and is denoted by $\operatorname{nrank}(A)$. Thus $\operatorname{nrank}(A) = \operatorname{rank}_{\mathbb{H}(\Omega)}(A) = \dim_{\mathbb{H}(\Omega)}(R(A))$.

Equivalently, $\operatorname{nrank}(A)$ is the number of linearly independent columns of A. Similarly, the dimension of the null space $N(A)$ is called the nullity of A, that is,

$$\operatorname{nullity}_{\mathbb{H}(\Omega)}(A) := \dim_{\mathbb{H}(\Omega)}(N(A)).$$

If $A \in \mathbb{H}(\Omega)^{m\times n}$ then it is easily seen that

$$A \text{ is regular } \iff \operatorname{nrank}(A) = m = n.$$

Hence A is singular when $m \neq n$ or when $\operatorname{nrank}(A) < \min(m, n)$.

Remark 2.1.7. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then there are two ranks of A, namely, the normal rank $\operatorname{nrank}(A)$ and the rank of the matrix $A(\omega)$ for each fixed $\omega \in \Omega$. It can be easily shown that $\operatorname{nrank}(A) = \max\{\operatorname{rank}(A(\omega)) : \omega \in \Omega\}$.*

Remark 2.1.8. *Let $B := \begin{bmatrix} f_1 & \cdots & f_\ell \end{bmatrix} \in \mathbb{H}(\Omega)^{n\times \ell}$. Then $\{f_1, \ldots, f_\ell\}$ is linearly independent $\iff \operatorname{nrank}(B) = \ell$. Hence B is an ordered basis of $R(B) \iff \operatorname{nrank}(B) = \ell$.*

We now consider analytic equivalence between holomorphic matrices.

Definition 2.1.9. *(a) Let $U \subset \Omega$ be open and $A, B \in \mathbb{H}(\Omega)^{m\times n}$. Then A and B are said to be equivalent on U if there exist $F \in \mathrm{GL}_m(\mathbb{H}(U))$ and $E \in \mathrm{GL}_n(\mathbb{H}(U))$ such that*

$$A = FBE, \quad \text{that is,} \quad A(z) = F(z)B(z)E(z) \text{ for all } z \in U.$$

We write $A \sim_U B$ or $A(z) \sim_U B(z)$ when A and B are equivalent on U.

(b) Let $\lambda \in \Omega$ and $A, B \in \mathbb{H}(\Omega)^{m\times n}$. Then A and B are said to be equivalent at λ if there exists an open neighbourhood U of λ such that $A \sim_U B$. We write $A \sim_\lambda B$ or $A(z) \sim_\lambda B(z)$ when A and B are equivalent at λ.

Remark 2.1.10. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Define the equivalent orbit* $\mathrm{Orb}_\Omega(A)$ *of A by*

$$\mathrm{Orb}_\Omega(A) := \{UAV : U \in \mathrm{GL}_m(\mathbb{H}(\Omega)) \text{ and } V \in \mathrm{GL}_n(\mathbb{H}(\Omega))\}.$$

Note that $A \sim_\Omega B \iff B \in \mathrm{Orb}_\Omega(A) \iff A \in \mathrm{Orb}_\Omega(B) \iff \mathrm{Orb}_\Omega(A) = \mathrm{Orb}_\Omega(B)$. Similarly, $A \sim_\lambda B \iff \mathrm{Orb}_U(A) = \mathrm{Orb}_U(B)$ for some open neighbourhood U of λ.

Observe that if $A \sim_\Omega B$ then $\mathrm{nrank}(B) = \mathrm{nrank}(A)$. Thus analytic equivalence preserves the normal rank of A. Hence $B \in \mathrm{Orb}_\Omega(A) \Longrightarrow \mathrm{nrank}(B) = \mathrm{nrank}(A)$.

Obviously, $A \sim_\Omega B \Longrightarrow A \sim_\lambda B$ for each $\lambda \in \Omega$. The following result, which is a special case of [38, Theorem 5.2], shows that the converse also holds.

Theorem 2.1.11. *Let $A, B \in \mathbb{H}(\Omega)^{m\times n}$. Then $A \sim_\Omega B \iff A \sim_\lambda B$ for each $\lambda \in \Omega$.*

We have defined the spectrum of $A \in \mathbb{H}(\Omega)^{n\times n}$ when A is regular. We now define the spectrum of an $m \times n$ holomorphic matrix which subsumes the notion of spectrum of a regular holomorphic matrix.

Definition 2.1.12 (spectrum)**.** *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then $\lambda \in \Omega$ is said to be an eigenvalue of A if $\mathrm{rank}(A(\lambda)) < \mathrm{nrank}(A)$. The spectrum $\sigma_\Omega(A)$ of A is given by*

$$\sigma_\Omega(A) := \{\lambda \in \Omega : \mathrm{rank}(A(\lambda)) < \mathrm{nrank}(A).\}$$

Thus the spectrum $\sigma_\Omega(A)$ consists of eigenvalues of A irrespective of whether A is regular or singular. Observe that $\sigma_\Omega(A)$ is invariant under analytic equivalence transformation of A on Ω. Hence

$$A \sim_\Omega B \Longrightarrow \sigma_\Omega(A) = \sigma_\Omega(B).$$

We now briefly review the Weierstrass product and the Weierstrass theorem for holomorphic functions which we utilize to prove Smith forms of holomorphic matrices.

2.1.1 Weierstrass factorization of holomorphic functions

Let $f \in \mathbb{H}(\Omega)$ and $\lambda \in \sigma_\Omega(f)$. A positive integer ℓ is said to be the order of λ as a zero of f if $f^{(\ell)}(\lambda) \neq 0$ and $f^{(j)}(\lambda) = 0$ for $j = 1 : \ell - 1$. Also, ℓ is called the multiplicity of λ as a zero of f. On the other hand, if $\mu \in \Omega$ and $f(\mu) \neq 0$ then μ is said to be a zero of f of order 0. Thus μ is a zero of f of order $0 \iff \mu \in \rho_\Omega(f)$.

We denote the set of integers by $\mathbb{Z}$ and the set of non-negative integers by $\mathbb{Z}_+$.

Definition 2.1.13. *Define* $o : \Omega \times (\mathbb{H}(\Omega) \setminus \{0\}) \longrightarrow \mathbb{Z}_+, (z, f) \longmapsto o_z(f)$, *by* $o_z(f) := p$, *where* p *is the order of* z *as a zero of* f.

Let $\phi : \Omega \longrightarrow \mathbb{C}$. The support of ϕ denoted by $\text{supp}(\phi)$ is given by

$$\text{supp}(\phi) := \{z \in \Omega : \phi(z) \neq 0\}.$$

The support of ϕ is said to be **locally finite** if $\text{supp}(\phi) \cap \text{K}$ is a finite set for every compact set $\text{K} \subset \Omega$.

Definition 2.1.14 (divisor)**.** *A function* $\partial : \Omega \longrightarrow \mathbb{Z}_+$ *is said to be a divisor on* Ω *if* $\text{supp}(\partial)$ *is locally finite. If* ∂_1 *and* ∂_2 *are divisors on* Ω *then we write* $\partial_1 \leq \partial_2$ *when* $\partial_1(z) \leq \partial_2(z)$ *for all* $z \in \Omega$. *We denote the set of all divisors on* Ω *by* $\text{Div}(\Omega)$.

Observe that if $\partial_1, \partial_2 \in \text{Div}(\Omega)$ then $\partial_1 + \partial_2 \in \text{Div}(\Omega)$. Also $\partial_{\max} := \max\{\partial_1, \partial_2\}$ and $\partial_{\min} := \min\{\partial_1, \partial_2\}$ are divisors on Ω, where

$$\partial_{\max}(z) := \max\{\partial_1(z), \partial_2(z)\} \text{ and } \partial_{\min}(z) := \min\{\partial_1(z), \partial_2(z)\} \text{ for } z \in \Omega.$$

Let $f \in \mathbb{H}(\Omega)$ be nonzero. Then $\sigma_\Omega(f)$ is a locally finite subset of Ω. Define $\partial : \Omega \longrightarrow \mathbb{Z}_+$ by $\partial(z) := o_z(f)$. Then $\text{supp}(\partial) = \sigma_\Omega(f)$ is locally finite. Hence $\partial \in \text{Div}(\Omega)$. Since $\partial(z) = o_z(f)$ is the order of z as a zero of f, it follows that $(z-\lambda)^{\partial(\lambda)}$ is a divisor of $f(z)$ for $\lambda \in \sigma_\Omega(f)$. Thus, multiplicities of zeros of f give a divisor on Ω.

Definition 2.1.15. *Let* $\partial \in \text{Div}(\Omega)$ *and* $f \in \mathbb{H}(\Omega)$ *be nonzero. Then* ∂ *is said to be the principal divisor of* f *on* Ω *if* $\partial(z) = o_z(f)$ *for all* $z \in \Omega$.

Remark 2.1.16. *Let* $f \in \mathbb{H}(\Omega)$ *be nonzero and* ∂ *be the principal divisor of* f *on* Ω. *Then* $\sigma_\Omega(f) = \text{supp}(\partial)$. *Hence it follows that* $\partial = 0 \iff f$ *is a unit element of* $\mathbb{H}(\Omega)$.

Note that ∂ *is the principal divisor of* uf *on* Ω *for any unit element* u *of* $\mathbb{H}(\Omega)$. *As we shall see, the converse is also true. If* f *and* g *in* $\mathbb{H}(\Omega)$ *have the same principal divisor* ∂ *on* Ω *then* $g = uf$ *for some unit element* u *of* $\mathbb{H}(\Omega)$.

If ∂ is the principal divisor of f on Ω with finite support and $\text{supp}(\partial) = \{\lambda_1, \ldots, \lambda_\ell\}$ then $f(z) = \prod_{j=1}^{\ell}(z-\lambda_j)^{\partial(\lambda_j)}g(z)$ for some unit element $g \in \mathbb{H}(\Omega)$. In contrast, if $\text{supp}(\partial) = \{\lambda_n : n \in \mathbb{N}\}$ then an infinite product representation of f of the form $f(z) = \prod_{n=1}^{\infty}(z-\lambda_n)^{\partial(\lambda_n)}g(z)$ for some unit element $g \in \mathbb{H}(\Omega)$ may not exists. However, an infinite product representation of f can be obtained by replacing the factors

$(z-\lambda_n)^{\partial(\lambda_n)}, n \in \mathbb{N}$, with so called Weierstrass factors, which are defined as follows.

Definition 2.1.17. *[48] An infinite product* $\prod_{n=1}^{\infty} f_n$ *with* $f_n = 1 + g_n \in \mathbb{H}(\Omega)$ *is said to be* **normally convergent** *in* Ω *if the series* $\sum_{n=1}^{\infty} g_n$ *converges absolutely on every compact set* $\mathrm{K} \subset \Omega$.

Observe that if $f_1, \ldots, f_\ell \in \mathbb{H}(\Omega)$ are nonzero then

$$\sigma_\Omega\Big(\prod_{j=1}^{\ell} f_j\Big) = \cup_{j=1}^{\ell}\sigma_\Omega(f_j) \text{ and } o_\lambda\Big(\prod_{j=1}^{\ell} f_j\Big) = o_\lambda(f_1) + \cdots + o_\lambda(f_\ell) \text{ for } \lambda \in \Omega.$$

Now, let $f := \prod_{n=1}^{\infty} f_n, f_n \in \mathbb{H}(\Omega)$ and $f_n \neq 0$ for $n \in \mathbb{N}$, be a normally convergent product in Ω. Then it is known [48] that $f \in \mathbb{H}(\Omega)$ and $f \neq 0$. Further, for each $\lambda \in \Omega$ there exists N such that $f_n(\lambda) \neq 0$ for all $n \geq N$. Hence, for each $\lambda \in \Omega$ there exists $N \in \mathbb{N}$ such that

$$f = f_1 f_2 \cdots f_{N-1} F_N \text{ and } F_N(\lambda) \neq 0, \text{ where } F_N := \prod_{n=N}^{\infty} f_n \in \mathbb{H}(\Omega). \tag{2.2}$$

Consequently, we have $\sigma_\Omega(f) = \cup_{n=1}^{\infty}\sigma_\Omega(f_n)$ and $o_\lambda(f) = \sum_{n=1}^{\infty} o_\lambda(f_n)$ for $\lambda \in \Omega$.

Definition 2.1.18. *Let* $\partial \in \mathrm{Div}(\Omega)$ *and* $\mathrm{S} \subset \mathrm{supp}(\partial)$. *A sequence* (λ_n) *is said to be a* ∂*-sequence in* S *if* $\lambda_n \in \mathrm{S}$ *for all* $n \in \mathbb{N}$ *and each* $\lambda \in \mathrm{S}$ *appears exactly* $\partial(\lambda)$ *times in the sequence.*

We now define Weierstrass product corresponding to a divisor on Ω.

Definition 2.1.19 ([48],Weierstrass product)**.** *Let* $\partial \in \mathrm{Div}(\Omega)$ *and* (λ_n) *be a* ∂*-sequence in* $\mathrm{supp}(\partial) \setminus \{0\}$. *Then an infinite product*

$$W(z) := z^{\partial(0)} \prod_{n=1}^{\infty} f_n(z),\ f_n \in \mathbb{H}(\Omega), \tag{2.3}$$

is called a Weierstrass product in Ω *for the divisor* ∂ *if the following conditions hold:*

(a) *The Weierstrass factor* f_n *has no zeros in* $\Omega \setminus \{\lambda_n\}$ *and* $o_{\lambda_n}(f_n) = 1$ *for all* $n \in \mathbb{N}$.

(b) *The product* $\prod_{n=1}^{\infty} f_n$ *converges normally in* Ω.

The Weierstrass product W in Ω for a divisor $\partial \in \mathrm{Div}(\Omega)$ is unique up to a unit element of $\mathbb{H}(\Omega)$. Nevertheless, we refer to W as the Weierstrass product in Ω for the divisor ∂. Note that if W is the Weierstrass product in Ω for $\partial \in \mathrm{Div}(\Omega)$ then ∂ is the principal divisor of W on Ω, that is, $\sigma_\Omega(W) = \mathrm{supp}(\partial)$ and $\partial(\lambda) = o_\lambda(W)$ for $\lambda \in \Omega$.

Theorem 2.1.20 ([48],Weierstrass product theorem). *(a) Let $\partial \in \mathrm{Div}(\mathbb{C})$ and (λ_n) be a ∂-sequence in $\mathrm{supp}(\partial)\setminus\{0\}$. If $\mathrm{supp}(\partial)$ has no accumulation points in $\mathbb{C}$ then there is a Weierstrass product in $\mathbb{C}$ for ∂ given by*

$$W(z) := z^{\partial(0)} \prod_{n=1}^{\infty} f_n(z), \;\; f_n \in \mathbb{H}(\mathbb{C}),$$

where $f_n(z) := \left(1 - \dfrac{z}{\lambda_n}\right)\exp\left(\dfrac{z}{\lambda_n} + \dfrac{1}{2}\left(\dfrac{z}{\lambda_n}\right)^2 + \cdots + \dfrac{1}{n-1}\left(\dfrac{z}{\lambda_n}\right)^{n-1}\right).$

(b) Let $\partial \in \mathrm{Div}(\Omega)$ and (λ_n) be a ∂-sequence in $\mathrm{supp}(\partial)$. If $\mathrm{supp}(\partial)$ has no accumulation points in Ω then there is a Weierstrass product in Ω for ∂ given by

$$W(z) := \prod_{n=1}^{\infty} f_n(z), \;\; f_n \in \mathbb{H}(\Omega)),$$

where $f_n, n \in \mathbb{N}$, are Weierstrass factors.

The following version of Theorem 2.1.20(b) is also known as Weierstrass theorem for holomorphic functions with prescribed zeros including their orders and is deduced from the Mittag-Leffler theorem; see [20, p.358].

Weierstrass Theorem in $\mathbb{H}(\Omega)$: *Let $(\lambda_n) \subset \Omega$ be a sequence of distinct complex numbers with no accumulation points in Ω. Let $(\ell_n) \subset \mathbb{N}$ be a sequence of distinct natural numbers. Then there exists $f \in \mathbb{H}(\Omega)$ such that $\sigma_\Omega(f) = \{\lambda_n : n \in \mathbb{N}\}$ and $o_{\lambda_n}(f) = \ell_n$ for all $n \in \mathbb{N}$.*

Proof. Define $\partial : \Omega \longrightarrow \mathbb{Z}$ by $\partial(\lambda_n) := \ell_n$ for all $n \in \mathbb{N}$ and $\partial(z) := 0$ for $z \neq \lambda_n$. Then we have $\partial \in \mathrm{Div}(\Omega)$. Hence by Theorem 2.1.20(b) there exists a Weierstrass product $f \in \mathbb{H}(\Omega)$ for ∂ such that $\sigma_\Omega(f) = \{\lambda_n : n \in \mathbb{N}\}$ and $o_{\lambda_n}(f) = \ell_n$ for all $\in \mathbb{N}$. □

A consequence of the Weierstrass product (Theorem2.1.20) is that a holomorphic function $f \in \mathbb{H}(\Omega)$ can be factorized as $f = uW$, where u is a unit element of $\mathbb{H}(\Omega)$ and W is the Weierstrass product for the principal divisor ∂ of f.

Theorem 2.1.21 (factorization, [48]). *(a) Let f be a nonzero entire function and ∂ be the principal divisor of f on $\mathbb{C}$. Let (λ_n) be a ∂-sequence in $\mathrm{supp}(\partial)\setminus\{0\}$. Then f can be written in the form*

$$f(z) = e^{g(z)} z^{\partial(0)} \prod_{n=1}^{\infty}\left[\left(1 - \frac{z}{\lambda_n}\right)\exp\left(\frac{z}{\lambda_n} + \frac{1}{2}\left(\frac{z}{\lambda_n}\right)^2 + \cdots + \frac{1}{n-1}\left(\frac{z}{\lambda_n}\right)^{n-1}\right)\right],$$

where $g \in \mathbb{H}(\mathbb{C})$ *and* $z^{\partial(0)} \prod_{n=1}^{\infty}(\cdots)$ *is a (possibly empty) Weierstrass product in* $\mathbb{C}$ *for the divisor* ∂.

(b) Let $f \in \mathbb{H}(\Omega)$ *be nonzero and* ∂ *be the principal divisor of* f *on* Ω. *Let* (λ_n) *be a* ∂*-sequence in* $\operatorname{supp}(\partial)$. *Then* f *can be written in the form*

$$f = u \prod_{n=1}^{\infty} f_n, \; f_n \in \mathbb{H}(\Omega),$$

where $u \in \mathbb{H}(\Omega)$ *is a unit element and* $\prod_{n=1}^{\infty} f_n$ *is a (possibly empty) Weierstrass product in* Ω *for the divisor* ∂.

Let $f, g \in \mathbb{H}(\Omega)$. Then f is called a divisor of g if $g/f \in \mathbb{H}(\Omega)$, that is, there exist $h \in \mathbb{H}(\Omega)$ such that $g = fh$. A nonzero non-unit $u \in \mathbb{H}(\Omega)$ is called a prime of $\mathbb{H}(\Omega)$ if u divides fg for some $f, g \in \mathbb{H}(\Omega)$ then u divides f or u divides g. The functions $\phi_\lambda : z \mapsto (z - \lambda)$ for $\lambda \in \Omega$, up to unit factors, are the primes of $\mathbb{H}(\Omega)$. We have the following result which is a kind of prime factorization of f.

Theorem 2.1.22 (prime factorization)**.** *Let* $f \in \mathbb{H}(\Omega)$ *be nonzero and* ∂ *be the principal divisor of* f *on* Ω. *Let* $\sigma_\Omega(f) = \{z_n : n \in \mathbb{N}\}$. *Then* f *can be written as*

$$f(z) = u(z) \prod_{n=1}^{\infty} (z - z_n)^{\partial(z_n)} u_n(z) \;\; \textit{for all } z \in \Omega,$$

where $u, u_n \in \mathbb{H}(\Omega)$ *are unit elements,* $n \in \mathbb{N}$ *and* $\prod_{n=1}^{\infty}(z - z_n)^{\partial(z_n)} u_n(z)$ *is a (possibly empty) product in* Ω *for the divisor* ∂.

Proof. Let (λ_n) be a ∂-sequence in $\operatorname{supp}(\partial)$. Then by Theorem 2.1.21(b), we have

$$f = u \prod_{n=1}^{\infty} f_n, \; f_n \in \mathbb{H}(\Omega),$$

where $u \in \mathbb{H}(\Omega)$ is a unit element. Note that the Weierstrass factor f_n has no zeros in $\Omega \backslash \{\lambda_n\}$ and that $o_{\lambda_n}(f_n) = 1$ for all $n \in \mathbb{N}$. Hence $f_n(z)/(z - \lambda_n)$ is a unit element of $\mathbb{H}(\Omega)$ as it has a removable singularity at λ_n for all $n \in \mathbb{N}$. Thus $f_n(z) = (z - \lambda_n)u_n(z)$ for some unit element $u_n \in \mathbb{H}(\Omega)$ for all $n \in \mathbb{N}$. Next, observe that each $\lambda \in \sigma_\Omega(f)$ is a simple zero of exactly $\partial(\lambda)$ Weierstrass factors $f_n, n \in \mathbb{N}$. Further, as $\prod_{n=1}^{\infty} f_n$ converges normally, for every bijection $\tau : \mathbb{N} \longrightarrow \mathbb{N}$ the rearrangement $\prod_{n=1}^{\infty} f_{\tau(n)}$ converges normally to the same limit [48, p.8]. Consequently, for each $z_n \in \sigma_\Omega(f)$, rearranging the Weierstrass product if necessary, we can collect together $\partial(z_n)$ Weierstrass factors with zero at z_n yielding the product $(z - z_n)^{\partial(z_n)} u_n(z)$ for some unit element $u_n \in \mathbb{H}(\Omega)$. Since

$\sigma_\Omega(f) = \{z_n : n \in \mathbb{N}\}$, it follows that there exist unit elements $u_n \in \mathbb{H}(\Omega)$ for all $n \in \mathbb{N}$ such that

$$\prod_{n=1}^{\infty} f_n(z) = \prod_{n=1}^{\infty} (z - z_n)^{\partial(z_n)} u_n(z).$$

This completes the proof. □

2.1.2 Smith forms of holomorphic matrices

Let $\lambda \in \mathbb{C}$ and let $\mathbb{H}(\lambda)$ denote the set of all complex functions holomorphic at λ. Thus $f \in \mathbb{H}(\lambda) \iff$ there exists an open neighbourhood U of λ such that $f \in \mathbb{H}(U)$. Let $f \in \mathbb{H}(\Omega)$ be nonzero and $\lambda \in \mathbb{C}$. Set $\nu(\lambda) := o_\lambda(f)$. Then there exists $g \in \mathbb{H}(\lambda)$ such that

$$f(z) = (z - \lambda)^{\nu(\lambda)} g(z) \ \text{ and } \ g(\lambda) \neq 0. \tag{2.4}$$

Since $g(\lambda) \neq 0$, there exists an open neighbourhood U of λ such that g is a unit element of $\mathbb{H}(U)$. Hence the local factorization in (2.4) can be restated as

$$f(z) \sim_\lambda (z - \lambda)^{\nu(\lambda)}, \ \text{where } \nu(\lambda) := o_\lambda(f).$$

In other words, $f(z)$ and $(z - \lambda)^{\nu(\lambda)}$ are locally equivalent at λ. A generalization of this result to the case of a matrix-valued holomorphic function is called the local Smith form.

The local Smith form is considered in [38] and [9, p.69] for a meromorphic matrix-valued function and in [35, p.414] for a holomorphic matrix valued-function. For completeness, we state and prove the local Smith form which will be useful later in proving the global Smith form of a holomorphic matrix-valued function.

Theorem 2.1.23. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Suppose that $nrank(A) = r$. Let $\lambda \in \Omega$. Then there exist non-negative integers $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ such that*

$$A(z) \sim_\lambda \left[\begin{array}{ccc|c} (z-\lambda)^{\nu_1(\lambda)} & & & \\ & \ddots & & \\ & & (z-\lambda)^{\nu_r(\lambda)} & \\ \hline & & & 0_{m-r\times n-r} \end{array}\right] =: S_\lambda(z).$$

Proof. Let A be given by $A(z) = [a_{ij}(z)]$ for $z \in \Omega$. Define $\nu_1(\lambda) = \min\{o_\lambda(a_{ij})\}$ where the minimum is taken over all nonzero functions a_{ij}. Then $\nu_1(\lambda) \geq 0$. By permuting rows and columns, if necessary, assume that $\nu_1(\lambda) = o_\lambda(a_{11})$. Then by local factorization of a_{11} at λ we have $a_{11}(z) = (z - \lambda)^{\nu_1(\lambda)} f_{11}(z)$ with $f_{11}(\lambda) \neq 0$ and $f_{11} \in \mathbb{H}(\lambda)$. Thus, $a_{11}(z) \sim_\lambda (z - \lambda)^{\nu_1(\lambda)}$. Also, we have $a_{ij}(z) = (z - \lambda)^{\nu_1(\lambda)} f_{ij}(z)$ for some $f_{ij} \in \mathbb{H}(\lambda)$.

Multiplying the first row of A by f_{i1}/f_{11} and subtracting from the i-th row for $i = 2 : m$ and then multiplying the first column of A by f_{1j}/f_{11} and subtracting from the j-th column for $j = 2 : n$, we have

$$A(z) \sim_\lambda \left[\begin{array}{c|ccc} (z-\lambda)^{\nu_1(\lambda)} & & & \\ \hline & g_{22}(z) & \cdots & g_{2n}(z) \\ & \vdots & \ddots & \vdots \\ & g_{m2}(z) & \cdots & g_{mn}(z) \end{array}\right] \text{ for some } g_{ij} \in \mathbb{H}(\lambda).$$

Again, let $\nu_2(\lambda) := \min\{o_\lambda(g_{ij})\}$ where the minimum is taken over all nonzero functions g_{ij}. Assume that $\nu_2(\lambda) = o_\lambda(g_{22})$. Then $g_{22}(z) = (z-\lambda)^{\nu_2(\lambda)} h_{11}(z)$ with $h_{11}(\lambda) \neq 0$ and $g_{ij}(z) = (z-\lambda)^{\nu_2(\lambda)} h_{ij}(z)$ for some $h_{ij} \in \mathbb{H}(\lambda)$. By construction, we have $\nu_1(\lambda) \leq \nu_2(\lambda)$. Again, by row and column elimination we obtain

$$A(z) \sim_\lambda \left[\begin{array}{cc|ccc} (z-\lambda)^{\nu_1(\lambda)} & & & & \\ & (z-\lambda)^{\nu_2(\lambda)} & & & \\ \hline & & b_{33}(z) & \cdots & b_{3n}(z) \\ & & \vdots & \cdots & \vdots \\ & & b_{m3}(z) & \cdots & b_{mn}(z) \end{array}\right] \text{ for some } b_{ij} \in \mathbb{H}(\lambda).$$

Repeating the procedure, we obtain the desired result after r steps. □

We write a block diagonal matrix $\mathrm{diag}(A_1, \cdots, A_m)$ by $A_1 \oplus \cdots \oplus A_m$. Then

$$A(z) \sim_\lambda (z-\lambda)^{\nu_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\nu_r(\lambda)} \oplus 0_{m-r\times n-r} = S_\lambda(z).$$

The diagonal matrix $S_\lambda(z)$ is called the **local Smith form** of $A(z)$ at $\lambda \in \Omega$. Observe that

$$\lambda \in \sigma_\Omega(A) \iff \mathrm{rank}(S_\lambda(\lambda)) < r \iff \nu_r(\lambda) \neq 0.$$

In particular, if $\nu_j(\lambda) \neq 0$ then $\lambda \in \sigma_\Omega(A)$. Suppose that $\lambda \in \sigma_\Omega(A)$. Then there exists $0 \leq \ell < r$ such that $\nu_j(\lambda) \neq 0$ for $j = \ell+1, \ldots, r$. The positive integers $\nu_{\ell+1}(\lambda), \ldots, \nu_r(\lambda)$ are called the **partial multiplicities** of the eigenvalue λ. Further,

$$\nu(\lambda) := \nu_{\ell+1}(\lambda) + \cdots + \nu_r(\lambda)$$

is called the **algebraic multiplicity** of the eigenvalue λ.

Definition 2.1.24 (index)**.** *Let* $A \in \mathbb{H}(\Omega)^{m\times n}$*. Suppose that* $\mathrm{nrank}(A) = r$*. Let* $\lambda \in \Omega$*. Then a tuple* $\mathrm{ind}(\lambda, A) := (\nu_1(\lambda), \nu_2(\lambda), \ldots, \nu_r(\lambda)) \in \mathbb{Z}_+^r$ *with* $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ *is said to be the index of* $A(z)$ *at* λ *if*

$$A(z) \sim_\lambda (z-\lambda)^{\nu_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\nu_r(\lambda)} \oplus 0_{m-r\times n-r}.$$

We denote the zero tuple in $\mathbb{Z}_+^r$ by 0. Then it follows that

$$\lambda \in \sigma_\Omega(A) \Longleftrightarrow \mathrm{ind}(\lambda, A) \neq 0.$$

Let $f \in \mathbb{H}(\Omega)$ be nonzero with the principal divisor $\partial \in \mathrm{Div}(\Omega)$. Suppose that $\sigma_\Omega(f) = \{z_n : n \in \mathbb{N}\}$. Then by the factorization Theorem 2.1.22 we have

$$f(z) \sim_\Omega \prod_{n=1}^{\infty} (z-z_n)^{\partial(z_n)} u_n(z),$$

where $u_n \in \mathbb{H}(\Omega)$ is a unit element and $n \in \mathbb{N}$. We show that the global Smith form is a diagonal representation of a holomorphic matrix-valued function in which the diagonal elements are Weierstrass products.

We now state and prove the global Smith form of a holomorphic matrix-valued function using divisors on Ω which is akin to the Smith form of a matrix polynomial [32, 25, 56, 50].

Theorem 2.1.25 (Smith form)**.** *Let* $A \in \mathbb{H}(\Omega)^{m\times n}$*. Suppose that* $nrank(A) = r$ *and* $\sigma_\Omega(A) = \{z_\ell : \ell \in \mathbb{N}\}$*. Then there exist holomorphic functions* $\phi_1, \ldots, \phi_r$ *in* $\mathbb{H}(\Omega)$ *with principal divisors* $\partial_1, \ldots, \partial_r$ *on* Ω *such that* $\partial_1 \leq \cdots \leq \partial_r$ *and*

$$A(z) \sim_\Omega \left[\begin{array}{ccc|c} \phi_1(z) & & & \\ & \ddots & & \\ & & \phi_r(z) & \\ \hline & & & 0_{m-r\times n-r} \end{array}\right] =: S_A(z).$$

The functions $\phi_1, \ldots, \phi_r$ *are unique up to unit elements of* $\mathbb{H}(\Omega)$ *and are given by*

$$\phi_j(z) = u_j(z) \prod_{\ell=1}^{\infty} (z-z_\ell)^{\partial_j(z_\ell)} u_{j\ell}(z) \ \text{ for all } z \in \Omega \text{ and } j = 1:r,$$

where $u_j, u_{j\ell} \in \mathbb{H}(\Omega)$ *are unit elements,* $\ell \in \mathbb{N}$ *and* $\prod_{\ell=1}^{\infty}(z-z_\ell)^{\partial_j(z_\ell)} u_{j\ell}(z)$ *is a (possibly empty) product in* Ω *for the divisor* ∂_j*. Further,* ϕ_j *divides* ϕ_{j+1} *for* $j = 1 : r-1$*.*

Proof. Let $\lambda \in \Omega$ and $\text{ind}(\lambda, A) = (\nu_1(\lambda), \ldots, \nu_r(\lambda))$ be the index of $A(z)$ at λ. Then we have the local Smith form

$$A(z) \sim_\lambda (z-\lambda)^{\nu_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\nu_r(\lambda)} \oplus 0_{m-r \times n-r} = S_\lambda(z).$$

For $j = 1 : r$, define $\partial_j : \Omega \longrightarrow \mathbb{Z}$ by $\partial_j(\lambda) := \nu_j(\lambda)$ for all $\lambda \in \Omega$. Then we have $\partial_1 \leq \cdots \leq \partial_r$. Note that $\partial_j(\lambda) = \nu_j(\lambda) \neq 0 \Longrightarrow \lambda \in \sigma_\Omega(A) \Longrightarrow \text{supp}(\partial_j) \subset \sigma_\Omega(A)$. This shows that $\text{supp}(\partial_j)$ is locally finite. Hence $\partial_1, \ldots, \partial_r$ are divisors on Ω.

By Theorem 2.1.20(b), let $\phi_1, \ldots, \phi_r$ be Weierstrass products in Ω for $\partial_1, \ldots, \partial_r$, respectively. Then $\partial_1, \ldots, \partial_r$ are the principal divisors of $\phi_1, \ldots, \phi_r$ on Ω, that is, $\partial_j(\lambda) = o_\lambda(\phi_j)$ for $j = 1 : r$. Hence by local factorization $\phi_j(z) \sim_\lambda (z-\lambda)^{\partial_j(\lambda)}$ for $j = 1 : r$. This shows that for each $\lambda \in \Omega$ we have

$$A(z) \sim_\lambda (z-\lambda)^{\partial_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\partial_r(\lambda)} \oplus 0_{m-r \times n-r} \sim_\lambda \phi_1(z) \oplus \cdots \oplus \phi_r(z) \oplus 0_{m-r \times n-r}.$$

Hence by Theorem 2.2.14, we have $A(z) \sim_\Omega \phi_1(z) \oplus \cdots \oplus \phi_r(z) \oplus 0_{m-r \times n-r} = S_A(z)$.

Since $\lambda \in \sigma_\Omega(A) \Longleftrightarrow \text{ind}(\lambda, A) \neq 0$, it follows that $\sigma_\Omega(A) = \cup_{j=1}^{r} \text{supp}(\partial_j)$. Hence the factorization of ϕ_j follows from Theorem 2.1.22 for $j = 1 : r$. Since $\partial_1 \leq \cdots \leq \partial_r$, it follows that ϕ_j divides ϕ_{j+1}, that is, $\phi_{j+1}/\phi_j \in \mathbb{H}(\Omega)$ for $j = 1 : r-1$. This completes the proof. □

Remark 2.1.26. *The diagonal matrix $S_A(z)$ is called the Smith canonical form of $A(z)$ on Ω. The functions $\phi_1, \ldots, \phi_r$ are called the* **invariant functions (or factors)** *of $A(z)$ on Ω. If ϕ_j is not a unit element then it is called a non-unit invariant function or a non-unit invariant factor of $A(z)$ on Ω. Define $\phi_A(z) := \prod_{j=1}^{r} \phi_j(z)$ for $z \in \Omega$. Then ϕ_A is unique up to a unit element of $\mathbb{H}(\Omega)$ and is called the* **zero function** *of $A(z)$ on Ω. It follows that $\lambda \in \Omega$ is an eigenvalue value of $A(z) \Longleftrightarrow \phi(\lambda) = 0$. Hence $\sigma_\Omega(A) = \{\lambda \in \Omega : \phi_A(\lambda) = 0\}$.*

In view of Theorem 2.1.25, we have the following result.

Proposition 2.1.27. *Let $A \in \mathbb{H}(\Omega)^{m \times n}$. Let the Smith form of $A(z)$ be given by*

$$S_A(z) := \phi_1(z) \oplus \cdots \oplus \phi_r(z) \oplus 0_{m-r \times n-r}.$$

Then $\text{ind}(\lambda, A) := (\nu_1(\lambda), \ldots, \nu_r(\lambda)) \subset \mathbb{Z}_+^r$ with $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ is the index of $A(z)$ at $\lambda \Longleftrightarrow \nu_j(\lambda) = o_\lambda(\phi_j)$ for $j = 1 : r$.

Proof. Suppose that $\nu_j(\lambda) = o_\lambda(\phi_j)$ for $j = 1 : r$. Then $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ and $\phi_j(\lambda) \sim_\lambda (z-\lambda)^{\nu_j(\lambda)}$ for $j = 1 : r$. Consequently, we have

$$A(z) \sim_\Omega S_A(z) \sim_\lambda (z-\lambda)^{\nu_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\nu_r(\lambda)} \oplus 0_{m-r\times n-r}$$

which shows that $\text{ind}(\lambda, A) = (\nu_1(\lambda), \ldots, \nu_r(\lambda))$ is the index of $A(z)$ at λ.

Conversely, if $\text{ind}(\lambda, A) = (\nu_1(\lambda), \ldots, \nu_r(\lambda))$ is the index of $A(z)$ at λ then, as shown in the proof of Theorem 2.1.25, we have $\phi_j(z) \sim_\lambda (z-\lambda)^{\nu_j(\lambda)}$ which in turn implies that $\nu_j(\lambda) = o_\lambda(\phi_j)$ for $j = 1 : r$. This completes the proof. $\square$

Let $A \in \mathbb{H}(\Omega)^{m\times n}$. Then $A(z) \sim_\Omega S_A(z)$, where $S_A(z)$ is the Smith canonical form of $A(z)$. Hence there exist $F \in \text{GL}_m(\mathbb{H}(\Omega))$ and $E \in \text{GL}_n(\mathbb{H}(\Omega))$ such that

$$F(z)A(z)E(z) = S_A(z) \ \text{ for all } \ z \in \Omega.$$

If $\text{nrank}(A) = r$ then the last $n-r$ columns of $E(z)$ form a basis of $N(A)$ and the first r columns of $F(z)^{-1}$ form a basis of $R(A)$. Indeed, let $E(z) = \begin{bmatrix} v_1(z) & \cdots & v_n(z) \end{bmatrix}$ and $F(z)^{-1} = \begin{bmatrix} u_1(z) & \cdots & u_m(z) \end{bmatrix}$ be column partitions of $E(z)$ and $F(z)^{-1}$. Since $S_A(z) = \text{diag}(\phi_1(z), \cdots, \phi_r(z), 0_{m-r\times n-r})$, we have $Av_j = 0$ for $j = r+1 : n$. This shows that $N(A) = \text{span}_{\mathbb{H}(\Omega)}(v_{r+1}, \ldots, v_n)$ and $\{v_{r+1}, \ldots, v_n\}$ is a basis of $N(A)$. Also, $Av_j = \phi_j u_j$ for $j = 1 : r$ which shows that $R(A) = \text{span}_{\mathbb{H}(\Omega)}(u_1, \ldots, u_r)$ and $\{u_1, \ldots, u_r\}$ is a basis of $R(A)$.

2.2 Canonical forms of meromorphic matrix-valued functions

A function $f : \Omega \longrightarrow X$ is said to be meromorphic in Ω if f is holomorphic in Ω except for poles. We denote the set of all X-valued meromorphic functions on Ω by $\mathbb{M}(\Omega, X)$, that is,

$$\mathbb{M}(\Omega, X) := \{f : \Omega \longrightarrow X \mid f \text{ is meromorphic in } \Omega\}.$$

Again, for simplicity of notation, when $X = \mathbb{C}$ and $X = \mathbb{C}^{m\times n}$, we set

$$\mathbb{M}(\Omega) := \mathbb{M}(\Omega, \mathbb{C}), \ \mathbb{M}(\Omega)^n := \mathbb{M}(\Omega, \mathbb{C}^n) \text{ and } \mathbb{M}(\Omega)^{m\times n} := \mathbb{M}(\Omega, \mathbb{C}^{m\times n}).$$

Thus $\mathbb{M}(\Omega)$ is the set of meromorphic functions and $\mathbb{M}(\Omega)^{m\times n}$ is the set of $m \times n$ matrices whose entries are meromorphic functions. We refer to elements of $\mathbb{M}(\Omega)^n$ as meromorphic vectors and elements of $\mathbb{M}(\Omega)^{m\times n}$ as meromorphic matrices.

Let $M \in \mathbb{M}(\Omega)^{m\times n}$. We denote the set of poles of M by $\wp_\Omega(M)$, that is,

$$\wp_\Omega(M) := \{z \in \Omega : z \text{ is a pole of } M(z)\}.$$

If $f \in \mathbb{M}(\Omega)$ then λ is a pole of $f \iff \lambda$ is a zero of $1/f$. Hence $\wp_\Omega(f)$ is a locally finite subset of Ω. The multiplicity of $\lambda \in \wp(f)$ as a zero of $1/f$ is called the order of λ as a pole of f. We now define the order function $o_\lambda(f)$ for a meromorphic function f. This subsumes the order function defined in Definition 2.1.13.

Definition 2.2.1. *Define* $o : \Omega \times (\mathbb{M}(\Omega) \setminus \{0\}) \longrightarrow \mathbb{Z}, (z, f) \longmapsto o_z(f)$, *by*

$$o_z(f) := \begin{cases} p & \text{if } p \text{ is the order of } z \text{ as a zero of } f, \\ -p & \text{if } p \text{ is the order of } z \text{ as a pole of } f. \end{cases}$$

We now define the divisors on Ω for analyzing zeros and poles of meromorphic functions. This subsumes the divisors defined in Definition 2.1.14

Definition 2.2.2 (divisor)**.** *A function* $\partial : \Omega \longrightarrow \mathbb{Z}$ *is said to be a divisor on* Ω *if* $\operatorname{supp}(\partial)$ *is locally finite. If* ∂_1 *and* ∂_2 *are divisors on* Ω *then we write* $\partial_1 \leq \partial_2$ *when* $\partial_1(z) \leq \partial_2(z)$ *for all* $z \in \Omega$*. A divisor* ∂ *on* Ω *is said to be non-negative if* $\partial \geq 0$*, that is,* $\partial(z) \geq 0$ *for all* $z \in \Omega$*. We denote the set of all divisors on* Ω *by* $\operatorname{Div}(\Omega)$.

Let $f \in \mathbb{M}(\Omega)$ *be nonzero and* $\partial \in \operatorname{Div}(\Omega)$*. Then* ∂ *is said to the principal divisor of* f *on* Ω *if* $\partial(z) := o_z(f)$ *for* $z \in \Omega$.

Let $\partial \in \operatorname{Div}(\Omega)$. Then $\partial = \partial^+ - \partial^-$ and $\operatorname{supp}(\partial^+) \cap \operatorname{supp}(\partial^-) = \emptyset$, where ∂^+ and ∂^- are non-negative divisors on Ω given by

$$\partial^+(z) := \max(0, \partial(z)) \;\text{ and }\; \partial^-(z) := \max(0, -\partial(z)) \;\text{ for }\; z \in \Omega. \tag{2.5}$$

The set of divisors $\operatorname{Div}(\Omega)$ on Ω is an additive abelian group.

Remark 2.2.3. *By convention* 0 *is a divisor with empty support, that is,* $\operatorname{supp}(0) = \emptyset$*. Let* $f \in \mathbb{M}(\Omega)$ *be nonzero and* $\partial \in \operatorname{Div}(\Omega)$ *be the principal divisor of* f *on* Ω*. Then it follows that* $f \in \mathbb{H}(\Omega) \iff \partial \geq 0$.

Let $f_1, f_2 \in \mathbb{H}(\Omega)$ be nonzero. Then $f_1/f_2 \in \mathbb{M}(\Omega)$. Let ∂_1 and ∂_2, respectively, be the principal divisors of f_1 and f_2 on Ω. Since $o_z(f_1/f_2) = o_z(f_1) - o_z(f_2)$ for $z \in \Omega$, it follows that $\partial := \partial_1 - \partial_2$ is the principal divisor of f_1/f_2 on Ω. Thus

$$f_2 \text{ divides } f_1 \iff f_1/f_2 \in \mathbb{H}(\Omega) \iff \partial \geq 0 \iff \partial_2 \leq \partial_1.$$

On the other hand, let $g \in \mathbb{M}(\Omega)$ be nonzero and let $\partial \in \operatorname{Div}(\Omega)$ be the principal divisor of g. Then $\partial = \partial^+ - \partial^-$, where ∂^+ and ∂^- are the non-negative divisors on Ω given by

(2.5). Now, by Weierstrass Theorem 2.1.20, there exist $f_1, f_2 \in \mathbb{H}(\Omega)$ such that ∂^+ and ∂^-, respectively, are the principal divisors of f_1 and f_2. Since $\mathrm{supp}(\partial^+) \cap \mathrm{supp}(\partial^-) = \emptyset$, the functions f_1 and f_2 do not have common zeros. Further, the principal divisor of f_1/f_2 on Ω is given by $\partial^+ - \partial^- = \partial$. Thus g and f_1/f_2 have the same principal divisor ∂ on Ω. Hence $g = f_1/f_2$ up to a unit element of $\mathbb{H}(\Omega)$. Indeed, the principal divisor of $g/(f_1/f_2)$ is $\partial - \partial = 0$ which implies that $g/(f_1/f_2)$ has no zeros in Ω and hence a unit element of $\mathbb{H}(\Omega)$. Thus $g = uf_1/f_2$ for some unit element $u \in \mathbb{H}(\Omega)$. This shows that if $g \in \mathbb{M}(\Omega)$ is nonzero then there exist $f_1, f_2 \in \mathbb{H}(\Omega)$ such that f_1 and f_2 do not have common zeros and $g = f_1/f_2$. In other words, $\mathbb{M}(\Omega)$ is the quotient field of the integral domain $\mathbb{H}(\Omega)$.

Corollary 2.2.4. *The set $\mathbb{M}(\Omega)$ is the quotient field of the integral domain $\mathbb{H}(\Omega)$. The set of meromorphic vectors $\mathbb{M}(\Omega)^n$ and the set of meromorphic matrices $\mathbb{M}(\Omega)^{m\times n}$ are vector spaces over the field $\mathbb{M}(\Omega)$.*

Let $\mathrm{S} \subset \mathbb{H}(\Omega)\setminus\{0\}$ be nonempty. Then $f \in \mathbb{H}(\Omega)$ is called a common divisor of S if f divides every element of S. A common divisor f of S is called a greatest common divisor (gcd) of S if every common divisor of S is a divisor of f. Greatest common divisors, when exist, are uniquely determined up to unit elements of $\mathbb{H}(\Omega)$. If f is a gcd of S then we write $f = \gcd(\mathrm{S})$. If $1 = \gcd(\mathrm{S})$ then S is called relatively prime. Note that S is relatively prime $\Longleftrightarrow \cap_{f\in \mathrm{S}}\sigma_\Omega(f) = \emptyset$. Also, if $\mathrm{U} \subset \mathbb{H}(\Omega)$ is nonempty and $g := \gcd(\mathrm{U})$ then $\gcd(\mathrm{S}\cup\mathrm{U}) = \gcd(f, g)$. If $f, g \in \mathbb{H}(\Omega)$ and $\gcd(f, g) = 1$ then by Wedderburn lemma [48, p.136] there exist $a, b \in \mathbb{H}(\Omega)$ such that $af + bg = 1$ and a has no zeros in Ω. Also, if $f \in \mathbb{H}(\Omega)$ is a gcd of the functions $f_1, \ldots, f_n$ in $\mathbb{H}(\Omega)$ then there exist [48, p.138] functions $a_1, \ldots, a_n$ in $\mathbb{H}(\Omega)$ such that $f = a_1f_1 + \cdots + a_nf_n$. Consequently, the ideal $\langle f_1, \ldots, f_n\rangle$ in $\mathbb{H}(\Omega)$ generated by $f_1, \ldots, f_n$ is given by the principal ideal $\langle f\rangle$, where $f = \gcd(f_1, \ldots, f_n)$. In other words, every finitely generated ideal in $\mathbb{H}(\Omega)$ is a principal ideal. Similarly, lcm of S also exists and is unique up to a unit element of $\mathbb{H}(\Omega)$.

Let $\partial_1, \ldots, \partial_n$ be non-negative divisors in $\mathrm{Div}(\Omega)$. Define $\partial_{\min}, \partial_{\max} : \Omega \longrightarrow \mathbb{Z}$ by

$$\partial_{\min}(z) := \min\{\partial_j(z) : j = 1 : n\} \text{ and } \partial_{\max}(z) := \max\{\partial_j(z) : j = 1 : n\}.$$

Then $\partial_{\min}, \partial_{\max} \in \mathrm{Div}(\Omega)$. If $\partial_1, \ldots, \partial_n$ are the principal divisors of the functions $f_1, \ldots, f_n$ in $\mathbb{H}(\Omega)$ then by Theorem 2.1.22 every function $f \in \mathbb{H}(\Omega)$ with $\partial_{\min}$ as the principal divisor on Ω is a gcd of $\{f_1, \ldots, f_n\}$, that is, $f = \gcd(f_1, \ldots, f_n)$. In fact, every nonempty subset of $\mathbb{H}(\Omega) \setminus \{0\}$ has a gcd. In particular, gcd of any finite set $\{f_1, \ldots, f_n\} \subset \mathbb{H}(\Omega)$ exists and is a linear combination of $f_1, \ldots, f_n$. In other words,

$\mathbb{H}(\Omega)$ is a Bezout domain [48, p.138]. Similarly, by Theorem 2.1.22 every function $f \in \mathbb{H}(\Omega)$ with $\partial_{\max}$ as the principal divisor on Ω is a lcm of $\{f_1, \ldots, f_n\}$, that is, $f = \operatorname{lcm}(f_1, \ldots, f_n)$.

The discussion above proves the following well known result; see [48].

Theorem 2.2.5. *Let $h \in \mathbb{M}(\Omega)$ be nonzero. Then there exist f and g in $\mathbb{H}(\Omega)$ such that $h = f/g$ and $\gcd(f, g) = 1$.*

The discussion above also yields Weierstrass theorem for a meromorphic function which can be deduced from the Mittag-Leffler theorem; see [20].

Theorem 2.2.6 (Weierstrass)**.** *Let $\partial \in \operatorname{Div}(\Omega)$ be such that $\operatorname{supp}(\partial)$ has no accumulation points in Ω. Then there exists $f \in \mathbb{M}(\Omega)$ such that ∂ is the principal divisor of f on Ω. Further, $\sigma_\Omega(f) = \operatorname{supp}(\partial)$ and $o_\lambda(f) = \partial(\lambda)$ for $\lambda \in \Omega$.*

Proof. By (2.5), we have $\partial = \partial^+ - \partial^-$ and $\operatorname{supp}(\partial^+) \cap \operatorname{supp}(\partial^-) = \emptyset$, where $\partial^+ \geq 0$ and $\partial^- \geq 0$ are divisors on Ω. Hence by Weierstrass Theorem 2.1.20, there exist $f_1, f_2 \in \mathbb{H}(\Omega)$ such that ∂^+ and ∂^-, respectively, are the principal divisors of f_1 and f_2. Since $\operatorname{supp}(\partial^+) \cap \operatorname{supp}(\partial^-) = \emptyset$, the functions f_1 and f_2 do not have common zeros, that is, $\gcd(f_1, f_2) = 1$. Define $f := f_1/f_2$. Then $f \in \mathbb{M}(\Omega)$ and ∂ is the principal divisor of f. Hence the results follow. □

The following version of Theorem 2.2.6 is known as Weierstrass theorem for a meromorphic function with prescribed zeros and poles including their orders and is deduced from the Mittag-Leffler theorem in [20, p.358]. Here is an alternative proof.

Weierstrass Theorem in $\mathbb{M}(\Omega)$: *Let $(\lambda_n) \subset \Omega$ be a sequence of distinct complex numbers with no accumulation points in Ω. Let $(\ell_n) \subset \mathbb{Z}$ be a sequence of distinct integers. Then there exists $f \in \mathbb{M}(\Omega)$ such that $\sigma_\Omega(f) = \{\lambda_n : n \in \mathbb{N}\}$ and $o_{\lambda_n}(f) = \ell_n$ for all $n \in \mathbb{N}$.*

Proof. Define $\partial : \Omega \longrightarrow \mathbb{Z}$ by $\partial(\lambda_n) := \ell_n$ and $\partial(z) := 0$ when $z \neq \lambda_n$ for all $n \in \mathbb{N}$. Then we have $\partial \in \operatorname{Div}(\Omega)$. Hence by Theorem 2.2.6 we have the desired function $f \in \mathbb{M}(\Omega)$. □

2.2.1 Smith-McMillan forms of meromorphic matrices

Recall that $\mathbb{M}(\Omega)$ is a field and $\mathbb{M}(\Omega)^n$ is a vector space over the field $\mathbb{M}(\Omega)$. The span of a subset $\{f_1, \ldots, f_\ell\} \subset \mathbb{M}(\Omega)^n$ is given by

$$\mathrm{span}_{\mathbb{M}(\Omega)}(f_1, \ldots, f_\ell) := \{r_1 f_1 + \cdots + r_\ell f_\ell : \ r_1, \ldots, r_\ell \in \mathbb{M}(\Omega)\}.$$

A subset $\{f_1, \ldots, f_\ell\} \subset \mathbb{M}(\Omega)^n$ is said to be **linearly dependent** if there exist $r_1, \ldots, r_\ell$ in $\mathbb{M}(\Omega)$ not all are zero such that $r_1 f_1 + \cdots + r_\ell f_\ell = 0$. Thus $\{f_1, \ldots, f_\ell\}$ is said to be **linearly independent** if it is not linearly dependent. Let $\mathcal{V} \subset \mathbb{M}(\Omega)^n$ be a subspace. Then $\mathcal{B} \subset \mathbb{H}(\Omega)^n$ is said to be a **basis** of $\mathcal{V}$ if $\mathcal{B}$ is linearly independent and $\mathrm{span}_{\mathbb{M}(\Omega)}(\mathcal{B}) = \mathcal{V}$. The dimension $\dim_{\mathbb{M}(\Omega)}(\mathcal{V})$ of $\mathcal{V}$ is the number of elements of a basis $\mathcal{B} \subset \mathbb{M}(\Omega)^n$.

Let $S := \{f_1, \ldots, f_\ell\} \subset \mathbb{H}(\Omega)^n \subset \mathbb{M}(\Omega)^n$. Then $\mathrm{span}_{\mathbb{H}(\Omega)}(S)$ is a submodule of $\mathbb{H}(\Omega)^n$ and $\mathrm{span}_{\mathbb{M}(\Omega)}(S)$ is a subspace of $\mathbb{M}(\Omega)^n$. It is easy to see that S is linearly independent in the module $\mathbb{H}(\Omega)^n \iff S$ is linearly independent in the vector space $\mathbb{M}(\Omega)^n$.

Proposition 2.2.7. *Let $S := \{f_1, \ldots, f_\ell\} \subset \mathbb{H}(\Omega)^n \subset \mathbb{M}(\Omega)^n$. Then S is a basis of* $\mathrm{span}_{\mathbb{H}(\Omega)}(S) \iff S$ *is a basis of* $\mathrm{span}_{\mathbb{M}(\Omega)}(S)$. *Hence*

$$\dim_{\mathbb{H}(\Omega)}\left(\mathrm{span}_{\mathbb{H}(\Omega)}(S)\right) = \dim_{\mathbb{M}(\Omega)}\left(\mathrm{span}_{\mathbb{M}(\Omega)}(S)\right).$$

Proof. If S is linearly independent in $\mathbb{M}(\Omega)^n$ then it is immediate that S is linearly independent in $\mathbb{H}(\Omega)^n$. Conversely, suppose that S is linearly independent in $\mathbb{H}(\Omega)^n$. Now consider $r_1 f_1 + \cdots + r_\ell f_\ell = 0$ for $r_1, \ldots, r_\ell$ in $\mathbb{M}(\Omega)$. By Theorem 2.2.5, we have $r_j = g_j/h_j$ and $\gcd(g_j, h_j) = 1$ where $g_j, h_j \in \mathbb{H}(\Omega)$ for $j = 1 : \ell$. Let $h := \mathrm{lcm}(h_1, \ldots, h_\ell)$. Then $hr_j \in \mathbb{H}(\Omega)$ for $j = 1 : \ell$. Hence $r_1 f_1 + \cdots + r_\ell f_\ell = 0 \implies hr_1 f_1 + \cdots + hr_\ell f_\ell = 0 \implies hr_1 = \cdots = hr_\ell = 0 \implies r_1 = \cdots = r_\ell = 0$. This shows that S is linearly independent in $\mathbb{M}(\Omega)^n$. Hence the results follow. □

Proposition 2.2.7 shows that if $S \subset \mathbb{H}(\Omega)^n$ then we can talk about linear independence of S without specifying $\mathbb{H}(\Omega)^n$ or $\mathbb{M}(\Omega)^n$.

Definition 2.2.8 (analytic basis). *Let $\mathcal{V}$ be subspace of $\mathbb{M}(\Omega)^n$ and $\mathcal{B} \subset \mathbb{H}(\Omega)^n$. Then $\mathcal{B}$ said to be an analytic basis of $\mathcal{V}$ if $\mathcal{B}$ is linearly independent and* $\mathrm{span}_{\mathbb{M}(\Omega)}(\mathcal{B}) = \mathcal{V}$.

Let $\mathcal{V} \subset \mathbb{M}(\Omega)^n$ be a subspace. If $\mathcal{V} \neq \{0\}$ then it has an analytic basis. Indeed, let $\{f_1, \ldots, f_\ell\} \subset \mathbb{M}(\Omega)^n$ be a basis of $\mathcal{V}$. Then by Theorem 2.2.5 entries of each f_j can be written as fractions of relatively coprime functions in $\mathbb{H}(\Omega)$. Now, taking lcm of the denominators of the entries of f_j, each f_j can be written as $f_j = g_j/h_j$ where $g_j \in \mathbb{H}(\Omega)^n$ and $h_j \in \mathbb{H}(\Omega)$ for $j = 1 : n$. Obviously, we have

$$\mathrm{span}_{\mathbb{M}(\Omega)}(f_1, \ldots, f_\ell) = \mathrm{span}_{\mathbb{M}(\Omega)}(g_1, \ldots, g_\ell).$$

Now $r_1 g_1 + \cdots + r_\ell g_\ell = 0 \Longrightarrow r_1 h_1 f_1 + \cdots + r_\ell h_\ell f_\ell = 0 \Longrightarrow r_1 h_1 = \cdots = r_\ell h_\ell = 0 \Longrightarrow r_1 = \cdots = r_\ell = 0$ showing that $\{g_1, \ldots, g_\ell\}$ is linearly independent in $\mathbb{M}(\Omega)^n$. This shows that $\{g_1, \ldots, g_\ell\} \subset \mathbb{H}(\Omega)$ is a basis of $\mathcal{V}$ and $\dim_{\mathbb{M}(\Omega)}(\mathcal{V}) = \ell$. This proves the following result.

Theorem 2.2.9. *Let $\mathcal{V} \subset \mathbb{M}(\Omega)^n$ be a subspace. If $\mathcal{V} \neq \{0\}$ then $\mathcal{V}$ has an analytic basis.*

Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Then for each fixed $\lambda \notin \wp_\Omega(A)$, we can talk about the rank of the matrix $A(\lambda) \in \mathbb{C}^{m\times n}$ over the field $\mathbb{C}$, which we denote by $\mathrm{rank}(A(\lambda))$. We can also talk about the rank of the matrix A when viewed as a linear transformation

$$A : \mathbb{M}(\Omega)^n \longrightarrow \mathbb{M}(\Omega)^m, f \longmapsto Af,$$

on vector spaces over the field $\mathbb{M}(\Omega)$. Consider the null and range spaces

$$N(A) := \{f \in \mathbb{M}(\Omega)^n : Af = 0\} \subset \mathbb{M}(\Omega)^n \text{ and } R(A) := \{Af : f \in \mathbb{M}(\Omega)^n\} \subset \mathbb{M}(\Omega)^m.$$

The dimensions of the null space $N(A)$ and the range space $R(A)$ over the field $\mathbb{M}(\Omega)$ are called nullity and normal rank of A, respectively.

Definition 2.2.10 (normal rank)**.** *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Then the rank of A is given by*

$$\mathrm{rank}_{\mathbb{M}(\Omega)}(A) := \dim_{\mathbb{M}(\Omega)}(R(A)).$$

The rank $\mathrm{rank}_{\mathbb{M}(\Omega)}(A)$ *is also called the normal rank of A and is denoted by* $\mathrm{nrank}(A)$. *Thus* $\mathrm{nrank}(A) = \mathrm{rank}_{\mathbb{M}(\Omega)}(A) = \dim_{\mathbb{M}(\Omega)}(R(A))$.

Definition 2.2.11 (regular)**.** *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Then A is said to be regular if $m = n$ and $A(z)$ is invertible for some $z \in \Omega \setminus \wp(\mathbf{A})$.*

If $A \in \mathbb{M}(\Omega)^{n\times n}$ then it is easily seen that

$$A \text{ is regular } \Longleftrightarrow \mathrm{nrank}(A) = n.$$

Thus $A \in \mathbb{M}(\Omega)^{m\times n}$ is singular when $m \neq n$ or when $\mathrm{nrank}(A) < \min(m, n)$.

Remark 2.2.12. *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Then there are two ranks of A, namely, the normal rank* $\text{nrank}(A)$ *and the rank of the matrix $A(\lambda)$ for each fixed $\lambda \notin \wp_\Omega(A)$. It can be easily shown that* $\text{nrank}(A) = \max\{\text{rank}(A(\lambda)) : \lambda \notin \wp_\Omega(A)\}$.

We now define equivalence between two meromorphic matrices.

Definition 2.2.13. *(a) Let $U \subset \Omega$ be open and $A, B \in \mathbb{M}(\Omega)^{m\times n}$. Then A and B are said to be equivalent on U if there exist $F \in \text{GL}_m(\mathbb{H}(U))$ and $E \in \text{GL}_n(\mathbb{H}(U))$ such that*

$$A = FBE, \quad \textit{that is,} \quad A(z) = F(z)B(z)E(z) \textit{ for all } z \in U.$$

We write $A \sim_U B$ or $A(z) \sim_U B(z)$ when A and B are equivalent on U.

(b) Let $\lambda \in \Omega$ and $A, B \in \mathbb{M}(\Omega)^{m\times n}$. Then A and B are said to be equivalent at λ if there exists an open neighbourhood U of λ such that $A \sim_U B$. We write $A \sim_\lambda B$ or $A(z) \sim_\lambda B(z)$ when A and B are equivalent at λ.

Observe that if $A \sim_\Omega B$ then $\text{nrank}(B) = \text{nrank}(A)$. Thus analytic equivalence preserves the normal rank of A. Obviously, $A \sim_\Omega B \Longrightarrow A \sim_\lambda B$ for each $\lambda \in \Omega$. The following result which is a special case of [38, Theorem 5.2] shows that the converse also holds.

Theorem 2.2.14. *Let $A, B \in \mathbb{M}(\Omega)^{m\times n}$. Then $A \sim_\Omega B \Longleftrightarrow A \sim_\lambda B$ for each $\lambda \in \Omega$.*

We have defined eigenvalues and spectrum of an $m \times n$ holomorphic matrix. As in the case of holomorphic matrices, the eigenvalues $A \in \mathbb{M}(\Omega)^{m\times n}$ can be defined in terms of the normal rank of A. However, the notion of spectrum of A requires appropriate modification.

Definition 2.2.15. *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Then $\lambda \in \Omega$ is said to be an eigenvalue of $A(z)$ if $\lambda \notin \wp_\Omega(A)$ and* $\text{rank}(A(\lambda)) < \text{nrank}(A)$. *The eigenspecturm $\text{eig}_\Omega(A)$ of A is given by*

$$\text{eig}_\Omega(A) := \{\lambda \in \Omega : \lambda \notin \wp_\Omega(A) \textit{ and } \text{rank}(A(\lambda)) < \text{nrank}(A)\}.$$

Let $\mu \in \wp_\Omega(A)$. Then μ said to be an eigenpole of A if there exists $f \in \mathbb{H}(\Omega)^n$ such that $f(\mu) \neq 0, f \notin N(A)$ and $\lim_{z\to\mu} A(z)f(z) = 0$. We denote the set of eigenpoles of A by $\text{eip}_\Omega(A)$. Then the spectrum $\sigma_\Omega(A)$ of $A(z)$ in Ω is given by

$$\sigma_\Omega(A) := \text{eig}_\Omega(A) \cup \text{eip}_\Omega(A).$$

If $\lambda \in \sigma_\Omega(A)$ then λ is said to be a zero of $A(z)$ in Ω.

Note that if $A \in \mathbb{H}(\Omega)^{m\times n}$ then $\sigma_\Omega(A) = \mathrm{eig}_\Omega(A)$. However, if $A \in \mathbb{M}(\Omega)^{m\times n}$ then $\mathrm{eig}_\Omega(A) \subset \sigma_\Omega(A)$ and the inclusion may be proper.

Example 2.2.16. *Consider $\Omega := \{z \in \mathbb{C} : |z| < 1\}$ and $A \in \mathbb{M}(\Omega)^{2\times 2}$ given by*

$$A(z) := \begin{bmatrix} 1 & \frac{1}{\sin z} \\ 0 & 1 \end{bmatrix}, \quad z \in \Omega.$$

Then $\mathrm{eig}_\Omega(A) = \emptyset$ and $\wp_\Omega(A) = \{0\}$. Consider $f(z) := \begin{bmatrix} 1 & -\sin z \end{bmatrix}^\top$. Then $f(0) \neq 0$ and $f \notin N(A)$. Further, we have $\lim_{z\to 0} A(z)f(z) = 0$ which shows that 0 is an eigenpole of A. Hence $\sigma_\Omega(A) = \mathrm{eip}_\Omega(A) = \{0\}$. ■

Remark 2.2.17. *Let $M \in \mathbb{M}(\Omega)^{m\times n}$ be given by $M(z) := [a_{ij}(z)]$, $z \in \Omega$. Then by Theorem 2.2.5 there exist $f_{ij}, g_{ij} \in \mathbb{H}(\Omega)$ such that $\gcd(f_{ij}, g_{ij}) = 1$ and $a_{ij} = f_{ij}/g_{ij}$ for all $i = 1 : m$ and $j = 1 : n$. Let $g := \mathrm{lcm}\{g_{ij}, i = 1 : m, j = 1 : n\}$. Note that g is unique up to a unit element of $\mathbb{H}(\Omega)$. Then we have $M = [f_{ij}/g_{ij}] = [h_{ij}]/g$, where $h_{ij} \in \mathbb{H}(\Omega)$ for $i = 1 : m$ and $j = 1 : n$. It follows that λ is a pole of $M(z) \iff g(\lambda) = 0$. Hence we have $\wp_\Omega(M) = \sigma_\Omega(g)$. The function g is called the pole function of M.*

A canonical form of a meromorphic matrix-valued function $A \in \mathbb{M}(\Omega)^{m\times n}$ is derived in [38]. It is shown that $A(z) \sim_\Omega f_1(z) \oplus \cdots \oplus f_r(z) \oplus 0_{m-r\times n-r}$, where $r = \mathrm{nrank}(A)$ and $f_1, \ldots, f_r$ are in $\mathbb{M}(\Omega)$ such that f_j divides f_{j+1} for $j = 1 : r$.

The Smith-McMillan form of a rational matrix is a well-known canonical form which has wide applications [32, 56]. We now prove the Smith-McMillan form of a meromorphic matrix-valued function which is akin to the Smith-McMillan form of a rational matrix.

Consider $h \in \mathbb{M}(\Omega) \setminus \{0\}$. Then there exist $f, g \in \mathbb{H}(\Omega)$ such that $\gcd(f, g) = 1$ and $h = f/g$. Let ∂_1 and ∂_2 be the principal divisors of f and g on Ω, respectively. Then f and g can be written as Weierstrass products in Ω for the divisors ∂_1 and ∂_2. The Smith-McMillan form generalizes this fact to the case of meromorphic matrix-valued functions.

Theorem 2.2.18 (Smith-McMillan form)**.** *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Suppose that*

$$\sigma_\Omega(A) = \{\lambda_n : n \in \mathbb{N}\} \text{ and } \wp_\Omega(A) = \{\mu_n : n \in \mathbb{N}\}.$$

Also suppose that $nrank(A) = r$. Then the following hold:

(a) There exist holomorphic functions $\phi_1, \ldots, \phi_r$ in $\mathbb{H}(\Omega)$ with principal divisors $\partial_1, \ldots, \partial_r$ on Ω such that $\partial_1 \leq \cdots \leq \partial_r$.

(b) *There exist holomorphic functions* $\psi_1, \ldots, \psi_r$ *in* $\mathbb{H}(\Omega)$ *with principal divisors* $\delta_1, \ldots, \delta_r$ *on* Ω *such that* $\delta_1 \geq \cdots \geq \delta_r$.

(c) ϕ_j *and* ψ_j *are relatively prime, that is,* $\gcd(\phi_j, \psi_j) = 1$ *for* $j = 1 : r$ *and*

$$A(z) \sim_\Omega \left[\begin{array}{ccc|c} \phi_1(z)/\psi_1(z) & & & \\ & \ddots & & \\ & & \phi_r(z)/\psi_r(z) & \\ \hline & & & 0_{m-r \times n-r} \end{array}\right] =: \Sigma_A(z),$$

where $\phi_1, \ldots, \phi_r$ *and* $\psi_1, \ldots, \psi_r$ *are unique up to unit elements of* $\mathbb{H}(\Omega)$. *Further,* ϕ_j *divides* ϕ_{j+1} *and* ψ_{j+1} *divides* ψ_j *for* $j = 1 : r - 1$.

(d) *Furthermore,* $\phi_1, \cdots, \phi_r$ *and* $\psi_1, \ldots, \psi_r$ *are given by*

$$\phi_j(z) = u_j(z) \prod_{\ell=1}^{\infty} (z - \lambda_\ell)^{\partial_j(\lambda_\ell)} u_{j\ell}(z) \text{ and } \psi_j(z) = v_j(z) \prod_{\ell=1}^{\infty} (z - \mu_\ell)^{\delta_j(\mu_\ell)} v_{j\ell}(z)$$

for $j = 1 : r$, *where* $u_j, v_j, u_{j\ell}, v_{j\ell} \in \mathbb{H}(\Omega)$ *are unit elements and* $\ell \in \mathbb{N}$. *Here* $\prod_{\ell=1}^{\infty} (z - \lambda_\ell)^{\partial_j(\lambda_\ell)} u_{j\ell}(z)$ *and* $\prod_{\ell=1}^{\infty} (z - \mu_\ell)^{\delta_j(\mu_\ell)} v_{j\ell}(z)$ *are (possibly empty) products in* Ω *for the divisors* ∂_j *and* δ_j, *respectively, for* $j = 1 : r$.

Proof. Let A be given by $A(z) := [a_{ij}(z)]$, $z \in \Omega$. Then there exist $f_{ij}, g_{ij} \in \mathbb{H}(\Omega)$ such that $\gcd(f_{ij}, g_{ij}) = 1$ and $a_{ij} = f_{ij}/g_{ij}$ for all $i = 1 : m$ and $j = 1 : n$. Let $g := \operatorname{lcm}(g_{ij}, i = 1 : m, j = 1 : n)$. Note that g is unique up to a unit element of $\mathbb{H}(\Omega)$. Then $A = [f_{ij}/g_{ij}] = [h_{ij}]/g = H/g$, where $H := [h_{ij}]$ and $h_{ij} \in \mathbb{H}(\Omega)$ for $i = 1 : m$ and $j = 1 : n$. Note that $\lambda \in \wp_\Omega(A) \iff g(\lambda) = 0$. Hence we have $\wp_\Omega(A) = \sigma_\Omega(g)$. By Theorem 2.1.25, the Smith form $S_H(z) = \widehat{\phi}_1(z) \oplus \cdots \oplus \widehat{\phi}_r \oplus 0_{m-r \times n-r}$ of H yields

$$A(z) = H(z)/g(z) \sim_\Omega \left[\begin{array}{ccc|c} \widehat{\phi}_1(z)/g(z) & & & \\ & \ddots & & \\ & & \widehat{\phi}_r(z)/g(z) & \\ \hline & & & 0_{m-r \times n-r} \end{array}\right],$$

where $\widehat{\phi}_1, \ldots, \widehat{\phi}_r$ are invariant factors of $H(z)$. Let $\phi_j, \psi_j \in \mathbb{H}(\Omega)$ be such that

$$\gcd(\phi_j, \psi_j) = 1 \text{ and } \frac{\phi_j}{\psi_j} = \frac{\widehat{\phi}_j}{g} \text{ for } j = 1 : r.$$

Since $\widehat{\phi}_j(z)$ divides $\widehat{\phi}_{j+1}(z)$, that is, $\widehat{\phi}_{j+1}/\widehat{\phi}_j \in \mathbb{H}(\Omega)$ there exists $f_j \in \mathbb{H}(\Omega)$ such that $\widehat{\phi}_{j+1} = f_j\widehat{\phi}_j$ for $j = 1 : r$. Hence for $j = 1 : r-1$, we have

$$\frac{\widehat{\phi}_{j+1}}{g} = f_j\frac{\widehat{\phi}_j}{g} \Longrightarrow \frac{\phi_{j+1}}{\psi_{j+1}} = f_j\frac{\phi_j}{\psi_j} \Longrightarrow \frac{\phi_{j+1}}{\psi_{j+1}}\frac{\psi_j}{\phi_j} = f_j.$$

Since ϕ_j and ψ_j are relatively prime and ϕ_j divides ϕ_{j+1}, it follows that ψ_{j+1} divides ψ_j for $j = 1, ..., r-1$.

Let $\partial_1, \ldots, \partial_r$ be the principal divisors of $\phi_1, \ldots, \phi_r$ on Ω. Also, let $\delta_1, \ldots, \delta_r$ be the principal divisors of $\psi_1, \ldots, \psi_r$ on Ω. As ϕ_j divides ϕ_{j+1} and ψ_{j+1} divides ψ_j for $j = 1 : r-1$, we have $\partial_1 \leq \cdots \leq \partial_r$ and $\delta_1 \geq \cdots \geq \delta_r$.

Next, observe that $\sigma_\Omega(A) = \sigma_\Omega(\Sigma_A)$ and $\wp_\Omega(A) = \wp_\Omega(\Sigma_A)$. Consequently, we have $\delta_j(\lambda) \neq 0 \Longrightarrow \psi_j(\lambda) = 0 \Longrightarrow \lambda \in \wp_\Omega(A)$. This shows that $\mathrm{supp}(\delta_j) \subset \wp_\Omega(A)$ for $j = 1 : r$. Hence the factorization of ψ_j follows from Theorem 2.1.22. On the other hand, $\partial_j(\lambda) \neq 0 \Longrightarrow \phi_j(\lambda) = 0$. Now either $\psi_i(\lambda) \neq 0$ for all $i = 1 : r$ or $\psi_i(\lambda) = 0$ for some $i \neq j$. In the first case, we have $\lambda \in \mathrm{eig}_\Omega(\Sigma_A) = \mathrm{eig}_\Omega(A)$. In the second case, $\lambda \in \wp_\Omega(\Sigma_A)$. Now, considering $v(z) := e_j$ we have $v(\lambda) \neq 0, v \notin N(\Sigma_A)$ and $\lim_{z\to\lambda} \Sigma_A(z)v(z) = \lim_{z\to\lambda} \phi_j(z)/\psi_j(z)e_j = 0$ as $\psi_j(\lambda) \neq 0$ which shows that $\lambda \in \mathrm{eip}_\Omega(\Sigma_A) = \mathrm{eip}_\Omega(A) \subset \sigma_\Omega(A)$. Thus, in either case, $\partial_j(\lambda) \neq 0 \Longrightarrow \lambda \in \sigma_\Omega(A)$. This shows that $\mathrm{supp}(\partial_j) \subset \sigma_\Omega(A)$ for $j = 1 : r$. Hence the factorization of ϕ_j follows from Theorem 2.1.22. This completes the proof. □

Remark 2.2.19. *The diagonal matrix $\Sigma_A(z)$ is called the Smith-McMillan form of $A(z)$ on Ω. The functions $\phi_1, \ldots, \phi_r$ are called the invariant zero functions (or factors) of $A(z)$ and the functions $\psi_1, \ldots, \psi_r$ are called the invariant pole functions (or factors) of $A(z)$ on Ω. Define $\phi_A(z) := \prod_{j=1}^r \phi_j(z)$ and $\psi_A(z) := \prod_{j=1}^r \psi_j(z)$. The function $\phi_A(z)$ is called the zero function of $A(z)$ and the function $\psi_A(z)$ is called the pole function of $A(z)$ on Ω. It follows that $\wp_\Omega(A) = \{\mu \in \Omega : \psi_A(\mu) = 0\} = \sigma_\Omega(\psi_A)$. It also follows from the proof of Theorem 2.2.18 that $\sigma_\Omega(\phi_A) = \{\lambda \in \Omega : \phi_A(\lambda) = 0\} \subset \sigma_\Omega(A)$. In fact, as the next result shows, the equality holds, that is, $\sigma_\Omega(A) = \sigma_\Omega(\phi_A)$.*

Theorem 2.2.20. *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Let ϕ_A and ψ_A be the zero and pole functions of A on Ω. Then $\sigma_\Omega(A) = \sigma_\Omega(\phi_A) = \mathrm{eig}_\Omega(A) \cup \mathrm{eip}_\Omega(A)$ and $\wp_\Omega(A) = \sigma_\Omega(\psi_A)$. Further,*

$$\mathrm{eig}_\Omega(A) = \{\lambda \in \Omega \ : \ \phi_A(\lambda) = 0 \text{ and } \psi_A(\lambda) \neq 0\},$$

$$\mathrm{eip}_\Omega(A) = \{\lambda \in \Omega : \ \phi_A(\lambda) = 0 \text{ and } \psi_A(\lambda) = 0\}.$$

Proof. Let the Smith-McMillan form $\Sigma_A(z)$ of $A(z)$ be as in Theorem 2.2.18. Since $A(z) \sim_\Omega \Sigma_A(z)$, it follows that $\sigma_\Omega(A) = \sigma_\Omega(\Sigma_A)$ and $\wp_\Omega(A) = \wp_\Omega(\Sigma_A)$. Next, observe that $\lambda \in \wp_\Omega(\Sigma_A) \iff \psi_i(\lambda) = 0$ for some $i \iff \psi_A(\lambda) = 0$. This shows that $\wp_\Omega(A) = \sigma_\Omega(\psi_A)$.

It follows that $\lambda \in \mathrm{eig}_\Omega(\Sigma_A) \iff \phi_j(\lambda) = 0$ for some j and $\psi_i(\lambda) \neq 0$ for $i = 1 : r \iff \phi_A(\lambda) = 0$ and $\psi_A(\lambda) \neq 0$. This shows that

$$\mathrm{eig}_\Omega(A) = \mathrm{eig}_\Omega(\Sigma_A) = \{\lambda \in \Omega : \phi_A(\lambda) = 0 \text{ and } \psi_A(\lambda) \neq 0\}.$$

Now suppose that $\phi_A(\lambda) = 0$ and $\psi_A(\lambda) = 0$. Then $\phi_j(\lambda) = 0$ and $\psi_j(\lambda) \neq 0$ for some j. Note that $\lambda \in \wp_\Omega(\Sigma_A)$. Consider $v(z) := e_j$ for $z \in \Omega$, where e_j is the j-th column of I_n. Then $v \in \mathbb{H}(\Omega)^n$ and we have $v(\lambda) \neq 0, v \notin N(\Sigma_A)$ and

$$\lim_{z\to\lambda} \Sigma_A(z)v(z) = \lim_{z\to\lambda} \frac{\phi_j(z)}{\psi_j(z)} e_j = 0$$

which shows that $\lambda \in \mathrm{eip}_\Omega(\Sigma_A)$.

Conversely, let $\mu \in \mathrm{eip}(\Sigma_A)$. Then there exists $v \in \mathbb{H}(\Omega)^n$ such that $v(\mu) \neq 0$, $v \notin N(\Sigma_A)$ and $\lim\limits_{z\to\mu} \Sigma_A(z)v(z) = 0$. Let $v(z)$ be given by $v(z) := \begin{bmatrix} v_1(z) & \cdots & v_n(z) \end{bmatrix}^\top$ for $z \in \Omega$. Then $\lim\limits_{z\to\mu} \Sigma_A(z)v(z) = 0 \implies \lim_{z\to\mu} \phi_j(z)v_j(z)/\psi_j(z) = 0$ for $j = 1 : r$. As $v(\mu) \neq 0$ and $v \notin N(\Sigma_A)$, we have $v_i(\mu) \neq 0$ for some $i \leq r$. This shows that $\lim_{z\to\mu} \phi_i(z)/\psi_i(z) = \lim_{z\to\mu} \frac{1}{v_i(\mu)} \lim_{z\to\mu} \phi_i(z)v_i(z)/\psi_i(z) = 0$. Since ϕ_i and ψ_i are co-prime, we have $\psi_i(\mu) \neq 0$ and $\phi_i(\mu) = 0$. Hence we have $\phi_A(\mu) = 0$ and $\psi_A(\mu) = 0$. This proves that $\mathrm{eip}_\Omega(\Sigma_A) = \{\lambda \in \Omega : \phi_A(\lambda) = 0 \text{ and } \psi_A(\lambda) = 0\}$.

Finally, $\sigma_\Omega(A) = \sigma_\Omega(\Sigma_A) = \sigma_\Omega(\phi_A)$ follows from the fact that if $\phi_A(\lambda) = 0$ then either $\psi_A(\lambda) \neq 0$ or $\psi_A(\lambda) = 0$. □

Notice that the spectrum $\sigma_\Omega(A)$ of a meromorphic matrix $A \in \mathbb{M}(\Omega)^{m\times n}$ consists of eigenvalues and eigenpoles. If $\lambda \in \sigma_\Omega(A)$ then λ is called a ***zero*** of $A(z)$. Thus, a zero of $A(z)$ is either an eigenvalue or an eigenpole of $A(z)$. We now define partial multiplicities of zeros and poles of $A(z)$. Let

$$\Sigma_A(z) = \phi_1/\psi_1 \oplus \cdots \oplus \phi_r/\psi_r \oplus 0_{m-r\times n-r}$$

be the Smith-McMillan form of $A(z)$. Recall that ϕ_i divides ϕ_{i+1} and ψ_{i+1} divides ψ_i for $i = 1 : r-1$. Let $\phi_A(z)$ and $\psi_A(z)$, respectively, be the zero function and pole function of $A(z)$ on Ω. Let $\lambda \in \Omega$. Define $\nu_j(\lambda) := o_\lambda(\phi_j)$ and $\kappa_j(\lambda) := o_\lambda(\psi_j)$ for $j = 1 : r$. Then $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ and $\kappa_1(\mu) \geq \cdots \geq \kappa_r(\mu)$ are non-negative integers such that

$$\phi_j(z) \sim_\lambda (z-\lambda)^{\nu_j(\lambda)} \text{ and } \psi_j(z) \sim_\lambda (z-\mu)^{\kappa_j(\mu)} \text{ for } j = 1 : r.$$

Define $\tau_j(\lambda) := o_\lambda(\phi_j/\psi_j)$ for $j = 1 : r$. Then it follows that $\tau_j(\lambda) = \nu_j(\lambda) - \kappa_j(\lambda) \in \mathbb{Z}$ for $j = 1 : r$. Since ϕ_j and ψ_j are co-prime, either $\tau_j(\lambda) = \nu_j(\lambda)$ or $\tau_j(\lambda) = -\kappa_j(\lambda)$ for $j = 1 : r$. Further, we have $\tau_1(\lambda) \leq \cdots \leq \tau_r(\lambda)$ and

$$\phi_j(z)/\psi_j(z) \sim_\lambda (z-\lambda)^{\tau_j(\lambda)} \text{ for } \ j = 1 : r.$$

Hence it follows that the following statements are equivalent.

(a) $\lambda \in \mathrm{eig}_\Omega(A)$.

(b) $\phi_A(\lambda) = 0$ and $\psi_A(\lambda) \neq 0$.

(c) $\nu_r(\lambda) \neq 0$ and $\kappa_1(\lambda) = 0$.

(d) $\tau_j(\lambda) = \nu_j(\lambda)$ for $j = 1 : r$.

Let $\lambda \in \mathrm{eig}_\Omega(A)$. Then $\phi_A(\lambda) = 0$ and hence there exists $0 \leq \ell < r$ such that

$$0 = \nu_1(\lambda) = \cdots = \nu_\ell(\lambda) < \nu_{\ell+1}(\lambda) \leq \cdots \leq \nu_r(\lambda). \tag{2.6}$$

The positive integers $\nu_{\ell+1}(\lambda), \ldots, \nu_r(\lambda)$ are called the ***partial multiplicities*** of λ. Further, $\nu(\lambda) := \nu_1(\lambda) + \cdots + \nu_r(\lambda) = \nu_{\ell+1}(\lambda) + \cdots + \nu_r(\lambda)$ is called the ***algebraic multiplicity*** of λ and $\nu_r(\lambda)$ is called the ***ascent*** of λ. Note that the algebraic multiplicity $\nu(\lambda)$ of an eigenvalue λ of A is the multiplicity of λ as a zero of $\phi_A(z)$.

Next, it follows that the following statements are equivalent for poles of $A(z)$ which are not zeros of $A(z)$.

(a) $\lambda \in \wp_\Omega(A) \setminus \mathrm{eip}_\Omega(A)$.

(b) $\phi_A(\lambda) \neq 0$ and $\psi_A(\lambda) = 0$.

(c) $\nu_r(\lambda) = 0$ and $\kappa_1(\lambda) \neq 0$.

(d) $\tau_j(\lambda) = -\kappa_j(\lambda)$ for $j = 1 : r$.

Let $\lambda \in \wp_\Omega(A)$. Then $\psi_A(\lambda) = 0$ and hence there exists $1 \leq \ell \leq r$ such that

$$\kappa_1(\lambda) \geq \cdots \geq \kappa_\ell(\lambda) > \kappa_{\ell+1}(\lambda) = \cdots = \kappa_r(\lambda) = 0. \tag{2.7}$$

The positive integers $\kappa_1(\lambda), \ldots, \kappa_\ell(\lambda)$ are called the ***partial multiplicities*** of the pole λ of $A(z)$ and $\kappa_1(\lambda)$ is the order of the pole λ. Further, $\kappa(\lambda) := \kappa_1(\lambda) + \cdots + \kappa_\ell(\lambda)$ is the total multiplicity of λ as a zero of the pole function $\psi_A(z)$.

Finally, it follows that the following statements are equivalent.

(a) $\lambda \in \text{eip}_\Omega(A)$.

(b) $\phi_A(\lambda) = 0$ and $\psi_A(\lambda) = 0$.

(c) $\nu_r(\lambda) \neq 0$ and $\kappa_1(\lambda) \neq 0$.

(d) $\tau_1(\lambda) = -\kappa_1(\lambda) < 0$ and $\tau_r(\lambda) = \nu_r(\lambda) > 0$.

Let $\lambda \in \text{eip}_\Omega(A)$. Then $\phi_A(\lambda) = 0$ and $\psi_A(\lambda) = 0$. Hence there exist positive integers ℓ and p such that $1 \leq \ell < p \leq r$ such that

$$\tau_1(\lambda) \leq \cdots \leq \tau_\ell(\lambda) < 0 < \tau_p(\lambda) \leq \cdots \leq \tau_r(\lambda). \tag{2.8}$$

The nonzero integers $\tau_1(\lambda), \ldots, \tau_\ell(\lambda), \tau_p(\lambda), \ldots, \tau_r(\lambda)$ are called the ***partial multiplicities*** of the eigenpole λ. Further,

$$\kappa(\lambda) := -(\tau_1(\lambda) + \cdots + \tau_\ell(\lambda)) = \kappa_1(\lambda) + \cdots + \kappa_\ell(\lambda)$$

is the total multiplicity of λ as a zero of $\psi_A(z)$ and

$$\nu(\lambda) := \tau_p(\lambda) + \cdots + \tau_r(\lambda) = \nu_p(\lambda) + \cdots + \nu_r(\lambda)$$

is the total multiplicity of λ as a zero of $\phi_A(z)$. Obviously, $-\tau_1(\lambda)$ is the order of λ as a pole of $A(z)$ and $\tau_r(\lambda)$ is the ascent of λ as a zero of $A(z)$.

In view of the observations above, we now define three indices of $A(z)$ at $\lambda \in \Omega$.

Definition 2.2.21. *Let $A \in \mathbb{M}(\Omega)^{m\times n}$. Let the Smith-McMillan form $\Sigma_A(z)$ of $A(z)$ be given by $\Sigma_A(z) = \phi_1/\psi_1 \oplus \cdots \oplus \phi_r/\psi_r \oplus 0_{m-r\times n-r}$.*

(a) **Zero index:** *Let $\lambda \in \text{eig}_\Omega(A)$. A tuple $\text{ind}_e(\lambda, A) := (\nu_1(\lambda), \ldots, \nu_r(\lambda)) \in \mathbb{Z}_+^r$ with $\nu_1(\lambda) \leq \cdots \leq \nu_r(\lambda)$ is called the zero index of $A(z)$ at λ if $\nu_j(\lambda) = o_\lambda(\phi_j)$ for $j = 1 : r$.*

(b) **Pole index:** *Let $\mu \in \wp_\Omega(A)$. A tuple $\text{ind}_p(\mu, A) := (\kappa_1(\mu), \ldots, \kappa_r(\mu)) \in \mathbb{Z}_+^r$ with $\kappa_1(\mu) \geq \cdots \geq \kappa_r(\mu)$ is called the pole index of $A(z)$ at μ if $\kappa_j(\mu) = o_\mu(\psi_j)$ for $j = 1 : r$.*

(c) **Pole-zero index:** *Let $\mu \in \sigma_\Omega(A) \cup \wp_\Omega(A)$. Then a tuple*

$$\text{ind}(\mu, A) := (\tau_1(\mu), \ldots, \tau_r(\mu)) \in \mathbb{Z}^r$$

with $\tau_1(\mu) \leq \cdots \leq \tau_r(\mu)$ is said to be the pole-zero index of $A(z)$ at μ if $\tau_j(\mu) = o_\mu(\phi_j/\psi_j)$ for $j = 1 : r$.

Remark 2.2.22. *(a) Observe that the pole-zero index* $\mathrm{ind}(\lambda, A)$ *of* $A(z)$ *at* λ *is determined by the zero index* $\mathrm{ind}_e(\lambda, A)$ *and the pole index* $\mathrm{ind}_p(\lambda, A)$ *of* $A(z)$ *at* λ*. Indeed,* $\mathrm{ind}(\lambda, A) = \mathrm{ind}_e(\lambda, A) - \mathrm{ind}_p(\lambda, A) = (\nu_1(\lambda) - \kappa_1(\lambda), \ldots, \nu_r(\lambda) - \kappa_r(\lambda))$.

(b) Note that if $\lambda \in \mathrm{eip}_\Omega(A)$ *then* λ *is a zero as well as a pole of* $A(z)$*. Hence* $A(z)$ *has two indices at* λ*, namely, the zero index* $\mathrm{ind}_e(\lambda, A)$ *and the pole index* $\mathrm{ind}_p(\lambda, A)$ *which separately provide information about zero structure and pole structure of* $A(z)$ *at* λ*. On the other hand, the pole-zero index* $\mathrm{ind}(\lambda, A)$ *provides information about pole-zero structure of* $A(z)$ *at the eigenpole* λ*.*

(c) Let $\mathrm{ind}(\lambda, A) := (\tau_1(\lambda), \ldots, \tau_r(\lambda)) \in \mathbb{Z}^r$ *be the pole-zero index of* $A(z)$ *at* $\lambda \in \Omega$*. Let* $\Sigma_A(z) := \phi_1(z)/\psi_1(z) \oplus \cdots \oplus \phi_r(z)/\psi_r(z) \oplus 0_{m-r,n-r}$ *be the Smith-McMillan form of* $A(z)$*. Then* $\phi_j(z)/\psi_j(z) \sim_\lambda (z-\lambda)^{\tau_j(\lambda)}$ *for* $j = 1 : r$*. Consequently, we have* $A(z) \sim_\Omega \Sigma_A(z) \sim_\lambda (z-\lambda)^{\tau_1(\lambda)} \oplus \cdots \oplus (z-\lambda)^{\tau_r(\lambda)} \oplus 0_{m-r\times n-r} =: \Sigma_\lambda(z)$*. The diagonal matrix* $\Sigma_\lambda(z)$ *is called the local Smith form of* $A(z)$ *at* λ*.*

2.3 Matrix-fraction description (MFD)

Let $A \in \mathbb{H}(\Omega)^{m\times n}$ and $D \in \mathbb{H}(\Omega)^{n\times n}$. Then D is said to be a **right divisor** of A if there exists $Q \in \mathbb{H}(\Omega)^{m\times n}$ such that $A = QD$. Let $B \in \mathbb{H}(\Omega)^{p\times n}$. If D is a right divisor of $B \in \mathbb{H}(\Omega)^{p\times n}$ then D is said to be a ***common right divisor*** of A and B. Further, D is said to be a ***greatest common divisor (gcrd)*** of A and B if $N \in \mathbb{H}(\Omega)^{n\times n}$ is a common right divisor of A and B then N is a right divisor of D. Furthermore, A and B are said to be ***right coprime*** if every gcrd of A and B are invertible, that is, if D is a gcrd of A and B then $D \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. Left divisors, common left divisors, greatest common left divisors, and left coprime matrices are defined similarly.

We mention that a common right divisor of A and B can be derived from the Smith canonical form. Indeed, let $E(z)A(z)F(z) = S_A(z)$ be the Smith form of $A(z)$. Then setting $D(z) := F(z)^{-1}$ and $Q(z) := E(z)^{-1}S_A(z)$, we have $A(z) = Q(z)D(z)$ which shows that D is a right divisor of A. Next, let $S(z)$ be the Smith form of $\mathcal{W} := \begin{bmatrix} A(z) \\ B(z) \end{bmatrix}$. Then $\mathcal{W} = E(z)S(z)F(z)$, for some $E \in \mathrm{GL}_{m+p}(\mathbb{H}(\Omega))$ and $F \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. Let $Q_1(z)$ (resp., $Q_2(z)$) denote the first m (resp., last p) rows of $E(z)S(z)$. Set $D(z) := F(z)$ Then we have

$$A(z) = Q_1(z)D \;\text{ and }\; B(z) = Q_2(z)D(z)$$

which show that $D(z)$ is a common right divisor of $A(z)$ and $B(z)$. A gcrd of two holomorphic matrices can also be extracted from the Smith canonical form.

Let $S_A(z)$ be the Smith form of A. Then $A(z) = E(z)S_A(z)F(z)$ for some $E \in \mathrm{GL}_m(\mathbb{H}(\Omega))$ and $E \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. We refer to the decomposition $A(z) = E(z)S_A(z)F(z)$ as the Smith decomposition of $A(z)$. Now if $\mathrm{nrank}(A) = r$ then

$$S_A(z) = \begin{bmatrix} D(z) & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_r \\ 0 \end{bmatrix} \begin{bmatrix} D(z) & 0 \end{bmatrix},$$

where $D(z) := \mathrm{diag}(\phi_1(z), \ldots, \phi_r(z))$ and $\phi_1, \ldots, \phi_r$ are invariant functions of A. Consider the conformal row and column partitions $E = \begin{bmatrix} E_1 & E_2 \end{bmatrix}$ and $F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$. Then we have $A(z) = E_1(z)D(z)F_1(z)$, where E_1 is left invertible and F_1 is right invertible, that is, there exist $E_L \in \mathbb{H}(\Omega)^{r\times m}$ and $F_R \in \mathbb{H}(\Omega)^{r\times n}$ such that $E_L(z)E_1(z) = I_r$ and $F_1(z)F_R(z) = I_r$ for all $z \in \Omega$. Define $D_R(z) := D(z)F_1(z)$. Then we have a (non-unique) decomposition

$$A(z) = E(z) \begin{bmatrix} D_R(z) \\ 0 \end{bmatrix} = E_1(z)D_R(z), \tag{2.9}$$

where $E_1(z)$ is left invertible. We refer to $D_R(z)$ as a ***right-structure matrix*** of $A(z)$. Observe that the Smith forms of $A(z)$ and $D_R(z)$ have the same non-unit invariant functions. Consequently, we have $\sigma_\Omega(A) = \sigma_\Omega(D_R)$ and $\mathrm{ind}_e(\lambda, A) = \mathrm{ind}_e(\lambda, D_R)$ for all $\lambda \in \sigma_\Omega(A)$. Thus $A(z)$ and $D_R(z)$ have the same spectral structure, that is, $A(z)$ and $D_R(z)$ have the same eigenvalues including their partial multiplicities. This proves the following result.

Proposition 2.3.1. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$ be such that* $\mathrm{nrank}(A) = r$*. Then A can be decomposed as $A(z) = E(z)D_R(z)$, where $E \in \mathbb{H}(\Omega)^{m\times r}$ is left invertible and $D_R \in \mathbb{H}(\Omega)^{r\times n}$ is a right-structure matrix of $A(z)$. Further, the Smith forms of $A(z)$ and $D_R(z)$ have the same non-unit invariant functions. In particular, $\sigma_\Omega(A) = \sigma_\Omega(D_R)$ and*

$$\mathrm{ind}_e(\lambda, A) = \mathrm{ind}_e(\lambda, D_R) \ \ \textit{for all} \ \ \lambda \in \sigma_\Omega(A).$$

The next result extracts a gcrd of two holomorphic matrices.

Theorem 2.3.2. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$ and $B \in \mathbb{H}(\Omega)^{p\times n}$. Set $\mathcal{W}(z) := \begin{bmatrix} A(z) \\ B(z) \end{bmatrix}$. Suppose that* $\mathrm{nrank}(\mathcal{W}) = n$*. Let $D_R \in \mathbb{H}(\Omega)^{n\times n}$ be a right-structure matrix of $\mathcal{W}(z)$. Then $D_R(z)$ is a gcrd of $A(z)$ and $B(z)$.*

Proof. Note that there is a matrix $E \in \mathrm{GL}_{m+p}(\mathbb{H}(\Omega))$ such that $\mathcal{W}(z) = E(z)\begin{bmatrix} D_R(z) \\ 0 \end{bmatrix}$.

Indeed, consider the Smith decomposition $\mathcal{W}(z) = E(z)\begin{bmatrix} D(z) \\ 0 \end{bmatrix} F(z)$, where $D \in \mathbb{H}(\Omega)^{n\times n}$ is given by $D(z) := \mathrm{diag}(\phi_1(z), \ldots, \phi_n(z))$. Then setting $D_R(z) := D(z)F(z)$, we have the desired result.

Now consider the conformal partition $E = \begin{bmatrix} E_1 & G_1 \\ E_2 & G_2 \end{bmatrix}$, where $E_1 \in \mathbb{H}(\Omega)^{m\times n}$ and $E_2 \in \mathbb{H}(\Omega)^{p\times n}$. Then $\begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$ is left invertible and $\mathcal{W}(z) = \begin{bmatrix} E_1(z) \\ E_2(z) \end{bmatrix} D_R(z)$ which shows that $A(z) = E_1(z)D_R(z)$ and $B(z) = E_2(z)D_R(z)$. Hence $D_R(z)$ is a common right divisor of $A(z)$ and $B(z)$.

Next, consider the conformal partition $E^{-1} = \begin{bmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{bmatrix}$. Then $E^{-1}\mathcal{W} = \begin{bmatrix} D_R \\ 0 \end{bmatrix}$ yields $X_1(z)A(z) + X_2(z)B(z) = D_R(z)$. If $N(z)$ is a common right divisor of $A(z)$ and $B(z)$ then $A(z) = Q_1(z)N(z)$ and $B(z) = Q_2(z)N(z)$. Consequently, we have $D_R(z) = X_1(z)A(z) + X_2(z)B(z) = (X_1(z)Q_1(z) + X_2(z)Q_2(z))N(z)$ showing that $N(z)$ is a right divisor of $D_R(z)$. This proves that $D_R(z)$ is a gcrd of $A(z)$ and $B(z)$. □

Recall that $A(z)$ and $B(z)$ are said to be right coprime if every gcrd of $A(z)$ and $B(z)$ is invertible. The next result characterizes right coprime matrices.

Proposition 2.3.3. *Let $A \in \mathbb{H}(\Omega)^{m\times n}$ and $B \in \mathbb{H}(\Omega)^{p\times n}$. Then $A(z)$ and $B(z)$ are right coprime* $\Longleftrightarrow$ rank $\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = n$ *for all* $z \in \Omega$.

Proof. Suppose that $A(z)$ and $B(z)$ are right coprime. Let D be a gcrd of $A(z)$ and $B(z)$. Then $D \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. Further, there exist $Q_1 \in \mathbb{H}(\Omega)^{m\times n}$ and $Q_2 \in \mathbb{H}(\Omega)^{p\times n}$ such that $\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$ is left invertible and $\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} D$. Now rank $\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} =$ rank $\begin{bmatrix} Q_1(z) \\ Q_2(z) \end{bmatrix} = n$ for all $z \in \Omega$.

Conversely, suppose that rank $\begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = n$ for all $z \in \Omega$. Set $\mathcal{W} := \begin{bmatrix} A \\ B \end{bmatrix}$. Then rank $\mathcal{W}(z) = n$ for all $z \in \Omega \Longrightarrow \mathrm{nrank}(\mathcal{W}) = n$. Let D be a right-structure matrix of $\mathcal{W}(z)$. Then by Theorem 2.3.2, $D(z)$ is a gcrd of $A(z)$ and $B(z)$. By Proposition 2.3.1, we have $\sigma_\Omega(D) = \sigma_\Omega(\mathcal{W}) = \emptyset$ which shows that D is invertible, that is, $D \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. □

Further, we have following result.

Theorem 2.3.4. *Let* $A \in \mathbb{H}(\Omega)^{m\times n}$ *and* $B \in \mathbb{H}(\Omega)^{p\times n}$*. Set* $\mathcal{W} := \begin{bmatrix} A \\ B \end{bmatrix}$*. Then the following statements are equivalent.*

(a) $A(z)$ *and* $B(z)$ *are right coprime*

(b) $\sigma_\Omega(\mathcal{W}) = \emptyset$.

(c) There exists $E \in \mathrm{GL}_{m+p}(\mathbb{H}(\Omega))$ *such that* $E(z)\mathcal{W}(z) = \begin{bmatrix} I_n \\ 0 \end{bmatrix} = S_{\mathcal{W}}(z)$*, where* $S_{\mathcal{W}}(z)$ *is the Smith form of* $\mathcal{W}(z)$.

(d) There exist $X \in \mathbb{H}(\Omega)^{n\times m}$ *and* $Y \in \mathbb{H}(\Omega)^{n\times p}$ *such that* $X(z)A(z)+Y(z)B(z) = I_n$ *for all* $z \in \Omega$.

(e) There exist holomorphic matrices $C(z)$ *and* $D(z)$ *such that* $\begin{bmatrix} A & C \\ B & D \end{bmatrix} \in \mathrm{GL}_{m+p}(\mathbb{H}(\Omega))$.

Proof. The proof follows from Smith form of $\mathcal{W}(z)$ and similar arguments as those used in the proofs of Theorem 2.3.2 and Proposition 2.3.3. □

We now explore solution of matrix equations of the form

$$X(z)A(z) + Y(z)B(z) = C(z) \text{ or } \begin{bmatrix} X(z) & Y(z) \end{bmatrix} \begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = C(z),$$

where the matrices $A \in \mathbb{H}(\Omega)^{m\times n}, B \in \mathbb{H}(\Omega)^{p\times n}$ and $C \in \mathbb{H}(\Omega)^{r\times n}$ are given and the matrices $X \in \mathbb{H}(\Omega)^{r\times m}$ and $Y \in \mathbb{H}(\Omega)^{r\times n}$ are sought to be determined.

Theorem 2.3.5. *Let* $A \in \mathbb{H}(\Omega)^{m\times n}, B \in \mathbb{H}(\Omega)^{p\times n}$ *and* $C \in \mathbb{H}(\Omega)^{r\times n}$*. Then the matrix equation*

$$X(z)A(z) + Y(z)B(z) = \begin{bmatrix} X(z) & Y(z) \end{bmatrix} \begin{bmatrix} A(z) \\ B(z) \end{bmatrix} = C(z)$$

has a solution $\begin{bmatrix} X(z) & Y(z) \end{bmatrix} \iff$ *every gcrd of* $A(z)$ *and* $B(z)$ *is a right divisor of* $C(z)$.

In particular, if $A(z)$ *and* $B(z)$ *are right coprime, then the matrix equation always has a solution.*

Proof. Suppose that a solution $\begin{bmatrix} X(z) & Y(z) \end{bmatrix}$ exists. If $D(z)$ is a gcrd of $A(z)$ and $B(z)$ then it follows that $D(z)$ is a right divisor of $C(z)$.

Conversely, suppose that every gcrd of $A(z)$ and $B(z)$ is a right divisor of $C(z)$. Let $D(z)$ be a gcrd of $A(z)$ and $B(z)$. Then $C(z) = R(z)D(z)$. Set $\mathcal{W}(z) := \begin{bmatrix} A(z) \\ B(z) \end{bmatrix}$. Then by (2.9), we have

$$E(z)\mathcal{W}(z) = \begin{bmatrix} D(z) \\ 0 \end{bmatrix},$$

where E is invertible. Consider the conformal partition $E(z) = \begin{bmatrix} Q_1(z) & Q_2(z) \\ P_1(z) & P_2(z) \end{bmatrix}$. Then $Q_1(z)A(z) + Q_2(z)B(z) = D(z)$. Left multiplying on both side by $R(z)$, we have

$$R(z)Q_1(z)A(z) + R(z)Q_2(z)B(z) = R(z)D(z) = C(z).$$

Setting $X(z) := R(z)Q_1(z)$ and $Y(z) := R(z)Q_2(z)$, we have

$$X(z)A(z) + Y(z)B(z) = C(z).$$

Hence $\begin{bmatrix} X(z) & Y(z) \end{bmatrix}$ is a solution of the matrix equation.

Now, if $A(z)$ and $B(z)$ are right coprime then a gcrd $D(z)$ is invertible. Then $Q_1(z)A(z) + Q_2(z)B(z) = D(z) \implies D(z)^{-1}Q_1(z)A(z) + D(z)^{-1}Q_2(z)B(z) = I_n$. Left multiplying on both sides by $C(z)$, we have

$$C(z)D(z)^{-1}Q_1(z)A(z) + C(z)D(z)^{-1}Q_2(z)B(z) = C(z).$$

Setting $X(z) := C(z)D(z)^{-1}Q_1(z)$ and $Y(z) := C(z)D(z)^{-1}Q_2(z)$, we obtain a solution $\begin{bmatrix} X(z) & Y(z) \end{bmatrix}$ of the matrix equation. □

We have already seen that if $f \in \mathbb{M}(\Omega)$ then there exist $p, q \in \mathbb{H}(\Omega)$ such that p and q are coprime and $f = \frac{p}{q}$. A similar matrix-faction description (MFD) is possible for a matrix-valued function $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. In fact, $\mathbf{M}(z)$ can be represented as

$$\mathbf{M}(z) = N_R(z)D_R(z)^{-1} = D_L(z)^{-1}N_L(z),$$

where N_L, N_R, D_L, D_R are holomorphic and D_L and D_R are regular. We discuss only a right MFD as the case of a left MFD is similar.

Definition 2.3.6 (Right MFD)**.** *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. A right matrix-fraction description (MFD) of* $\mathbf{M}(z)$ *is a decomposition of* $\mathbf{M}(z)$ *of the form* $\mathbf{M}(z) = N(z)D(z)^{-1}$*, where* $N \in \mathbb{H}(\Omega)^{m\times n}$ *and* $D \in \mathbb{H}(\Omega)^{n\times n}$ *with* $D(z)$ *being regular.*

A right MFD $\mathbf{M}(z) = N(z)D(z)^{-1}$ *is a said to be right coprime if* $N(z)$ *and* $D(z)$ *are right coprime, that is, if* $\operatorname{rank}\begin{bmatrix} N(z) \\ D(z) \end{bmatrix} = n$ *for* $z \in \mathbb{C}$*.*

A right coprime MFD of $\mathbf{M}(z)$ exists and can be deduced from the Smith-McMillan form of $\mathbf{M}(z)$. Indeed, we have the following result.

Proposition 2.3.7. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. Then there exist non-unique pair* $(N, D) \in \mathbb{M}(\Omega)^{m\times n} \times \mathbb{M}(\Omega)^{n\times n}$ *such that* $N(z)$ *and* $D(z)$ *are right coprime,* $D(z)$ *is regular, and*

$$\mathbf{M}(z) = N(z)D(z)^{-1}.$$

Proof. Let $\Sigma_{\mathbf{M}}(z) = \operatorname{diag}(\phi_1(z)/\psi_1(z), \ldots, \phi_r(z)/\psi_r(z), 0_{m-r,n-r})$ be the Smith-McMillan form of $\mathbf{M}(z)$. Then by Theorem 2.2.18, we have $\mathbf{M}(z) = E(z)\Sigma_{\mathbf{M}}(z)F(z)$, where E and F are invertible. Define

$$N_\phi(z) := \operatorname{diag}(\phi_1(z), \cdots, \phi_r(z)) \oplus 0_{(m-r)\times(n-r)} \text{ and } D_\psi(z) := \operatorname{diag}(\psi_1(z), ..., \psi_r(z)) \oplus I_{n-r}.$$

Then $\Sigma_{\mathbf{M}}(z) = N_\phi(z)D_\psi(z)^{-1}$ is a right coprime MFD of $\Sigma_{\mathbf{M}}(z)$. Since E and F are invertible, setting $N(z) := E(z)N_\phi(z)$ and $D(z) := F(z)^{-1}D_\psi(z)$, it follows that $N(z)$ and $D(z)$ are right coprime and that $\mathbf{M}(z) = E(z)\Sigma_{\mathbf{M}}(z)F(z) = N(z)D(z)^{-1}$. Indeed,

$$\operatorname{rank}\begin{bmatrix} N(z) \\ D(z) \end{bmatrix} = \operatorname{rank}\begin{bmatrix} E(z)N_\phi(z) \\ F(z)^{-1}D_\psi(z) \end{bmatrix} = \operatorname{rank}\begin{bmatrix} N_\phi(z) \\ D_\psi(z) \end{bmatrix} = n$$

for all $z \in \Omega$. Hence $\mathbf{M}(z) = N(z)D(z)^{-1}$ is a right coprime MFD. □

We now show that a right coprime MFD of $\mathbf{M}(z)$ is unique up to a unit element.

Theorem 2.3.8. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. If* $\mathbf{M}(z) = N_1(z)D_1(z)^{-1} = N_2(z)D_2(z)^{-1}$ *are right coprime MFDs then there exists* $U \in \mathrm{GL}_n(\mathbb{H}(\Omega))$ *such that* $N_1(z) = N_2(z)U(z)$ *and* $D_1(z) = D_2(z)U(z)$*.*

In particular, if $D_1(z)$ *and* $D_2(z)$ *are matrix polynomials then* $U(z)$ *is a unimodular matrix polynomial.*

Proof. Note that $\wp_\Omega(\mathbf{M}) = \sigma_\Omega(D_1) = \sigma_\Omega(D_2)$. Since $N_1(z)D_1(z)^{-1} = N_2(z)D_2(z)^{-1}$, we have $N_1(z) = N_2(z)D_2(z)^{-1}D_1(z) = N_2(z)U(z)$, where $U(z) := D_2(z)^{-1}D_1(z)$. Then

$U \in \mathbb{M}(\Omega)^{n\times n}$. We show that $U \in \mathrm{GL}_n(\mathbb{H}(\Omega))$ by proving that $D_2(z)^{-1}D_1(z)$ and its inverse $D_1(z)^{-1}D_2(z)$ are both analytic in Ω.

Since $N_1(z)$ and $D_1(z)$ are right coprime, by Theorem 2.3.4(d), there exist analytic matrices $X \in \mathbb{H}(\Omega)^{n\times m}$ and $Y \in \mathbb{H}(\Omega)^{n\times n}$ such that $X(z)N_1(z) + Y(z)D_1(z) = I_n$. Inserting $N_1(z) = N_2(z)U(z)$ we have

$$X(z)N_2(z)U(z) + Y(z)D_1(z) = I_n \Longrightarrow [X(z)N_2(z) + Y(z)D_2(z)]U(z) = I_n.$$

This shows that $U(z)^{-1} = X(z)N_2(z) + Y(z)D_2(z)$ is analytic in Ω.

Similarly, considering $V(z) := D_1(z)^{-1}D_2(z)$ and interchanging the role of $D_1(z)$ and $D_2(z)$ in the above proof, it follows that $V(z)^{-1}$ is analytic in Ω. Since $V(z)^{-1} = U(z)$ and both $U(z)$ and $U(z)^{-1}$ are analytic in Ω, we have $U \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. This proves that $N_1(z) = N_2(z)U(z)$ and $D_1(z) = D_2(z)U(z)$.

Now, if $D_1(z)$ and $D_2(z)$ are matrix polynomials, then $U(z) = D_1(z)D_2(z)^{-1}$ is a rational matrix. Since both $U(z)$ and $U(z)^{-1}$ are analytic, both $U(z)$ and $U(z)^{-1}$ must be matrix polynomials. In other words, $U(z)$ is a unimodular matrix polynomial. □

In view of the proof of Proposition 2.3.7, we have the following result.

Theorem 2.3.9. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. *Let* $\mathbf{M}(z) = N(z)D(z)^{-1}$ *be a right coprime MFD. Let* $\Sigma_{\mathbf{M}}(z) := \mathrm{diag}(\phi_1(z)/\psi_1(z), \ldots, \phi_r(z)/\psi_r(z), 0_{m-r,n-r})$ *be the Smith-McMillan form of* $\mathbf{M}(z)$. *Then* $S_N(z) := \mathrm{diag}(\phi_1(z), \ldots, \phi_r(z), 0_{m-r,n-r})$ *is the Smith form of* $N(z)$ *and* $S_D(z) := I_{n-r} \oplus \mathrm{diag}(\psi_r(z), \psi_{r-1}(z), \ldots, \psi_1(z))$ *is the Smith form of* $D(z)$.

In particular, we have

$$\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(N), \quad \mathrm{Ind}_e(\lambda, \mathbf{M}) = \mathrm{Ind}_e(\lambda, N) \ \ \textit{for} \ \ \lambda \in \sigma_\Omega(\mathbf{M})$$

$$\wp_\Omega(\mathbf{M}) = \sigma_\Omega(D), \quad \mathrm{Ind}_p(\lambda, \mathbf{M}) = \mathrm{Ind}_e(\lambda, D) \ \ \textit{for} \ \ \lambda \in \wp_\Omega(\mathbf{M}).$$

Proof. The proof of Proposition 2.3.7 shows that

$$\mathbf{M}(z) = E(z)\Sigma_{\mathbf{M}}(z)F(z) = E(z)N_\phi(z)D_\psi(z)^{-1}F(z) = N(z)D(z)^{-1}$$

are right coprime MFDs. Hence by Theorem 2.3.8, there exists $U \in \mathrm{GL}_n(\mathbb{H}(\Omega))$ such that $N(z) = E(z)\mathbf{N}_\phi(z)U(z)$ and $D(z) = F(z)^{-1}D_\psi(z)U(z)$. Consequently,

$$S_N(z) = S_{N_\phi}(z) = N_\phi(z) = \mathrm{diag}(\phi_1(z), \ldots, \phi_r(z), 0_{m-r,n-r})$$

is the Smith form of $N(z)$ and

$$S_D(z) = S_{D_\psi}(z) = \mathrm{diag}(I_{n-r}, \psi_r(z), \psi_{r-1}(z), \ldots, \psi_1(z))$$

is the Smith form of $D(z)$.

Obviously, the zeros and poles of $\mathbf{M}(z)$ are given by the eigenvalues of $N_\phi(z)$ and $D_\psi(z)$, respectively. In fact, we have

$$\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(N_\phi) = \sigma_\Omega(N) \text{ and } \wp_\Omega(\mathbf{M}) = \sigma_\Omega(D_\psi) = \sigma_\Omega(D).$$

Further, we have

$$\begin{aligned} \mathrm{Ind}_e(\lambda, \mathbf{M}) &= \mathrm{Ind}_e(\lambda, N_\phi) = \mathrm{Ind}_e(\lambda, N) \text{ for } \lambda \in \sigma(\mathbf{M}) \\ \mathrm{Ind}_p(\lambda, \mathbf{M}) &= \mathrm{Ind}_e(\lambda, D_\psi) = \mathrm{Ind}_e(\lambda, D) \text{ for } \lambda \in \wp_\Omega(\mathbf{M}). \end{aligned}$$

□

We mention that similar results hold for left coprime MFDs of $\mathbf{M}(z)$. We end this section with a proof of the argument principle for a regular meromorphic matrix-valued function using MFD.

The argument principle is a well known classical result in complex analysis [58, p.282] which is stated as follows.

Let $f : \Omega \to \mathbb{C}$ be a meromorphic and $\Gamma \subset \Omega$ be a rectifiable simple closed curve such that $\mathrm{Int}(\Gamma) \subset \Omega$. Here $\mathrm{Int}(\Gamma)$ denotes the interior of the region enclosed by Γ. Suppose that f is analytic and nonzero on Γ. Let P be the number of poles and N be the number of zeros of f inside Γ counting multiplicities. Then [58, p.282]

$$\frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz = N - P.$$

We now use MFD to provide a proof of the argument principle for a meromorphic matrix-valued function.

Theorem 2.3.10 (argument principle)**.** *Let* $\mathbf{M} \in \mathbb{M}(\Omega, \mathbb{C}^{n\times n})$ *be regular. Let* $\Gamma \subset \Omega$ *be a rectifiable simple closed curve such that* $\mathrm{Int}(\Gamma) \subset \Omega$*. Suppose that* $\mathbf{M}$ *is analytic on* Γ *and that* Γ *does not pass through an eigenvalue of* $\mathbf{M}(z)$*. Let* P *be the number of poles (counting multiplicities) and* N *be the number zeroes (counting multiplicities) of* $\mathbf{M}(z)$ *inside* Γ*. Then*

$$\frac{1}{2\pi i}\int_{\partial\Omega} \mathrm{Tr}\left(\mathbf{M}(z)^{-1}\mathbf{M}(z)'\right) dz = N - P.$$

Proof. Let $\mathbf{M}(z) = N(z)D(z)^{-1}$ be a right coprime MFD. By Theorem 2.3.9, the number of eigenvalues of $N(z)$ inside Γ = the number of zeroes of $\mathbf{M}(z)$ inside $\Gamma = N$, and the number of eigenvalues of $D(z)$ inside Γ = the number of poles of $\mathbf{M}(z)$ inside $\Gamma = P$.

Differentiating $\mathbf{M}(z)$ with respect to z, we have

$$\begin{aligned}\mathbf{M}(z)' &= N(z)'D(z)^{-1} + N(z)(D(z)^{-1})' \\ &= N(z)'D(z)^{-1} + N(z)\left(-D(z)^{-1}D(z)'D(z)^{-1}\right).\end{aligned}$$

Hence

$$\begin{aligned}\mathrm{Tr}\left((\mathbf{M}(z)^{-1}\mathbf{M}(z)'\right)dz &= \mathrm{Tr}\left(D(z)N(z)^{-1}.\left(N(z)'D(z)^{-1} - N(z)D(z)^{-1}D(z)'D(z)^{-1}\right)\right) \\ &= \mathrm{Tr}\left(D(z)N(z)^{-1}N(z)'D(z)^{-1}\right) - \mathrm{Tr}\left(D(z)'(D(z)^{-1})\right) \\ &= \mathrm{Tr}\left(N(z)^{-1}N(z)'\right) - \mathrm{Tr}\left(D(z)'(D(z)^{-1})\right).\end{aligned}$$

Consequently, we have

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial\Omega}\mathrm{Tr}\left((\mathbf{M}(z)^{-1}\mathbf{M}(z)'\right)dz &= \frac{1}{2\pi i}\int_{\partial\Omega}\mathrm{Tr}\left(N(z)^{-1}N(z)'\right) - \frac{1}{2\pi i}\int_{\partial\Omega}\mathrm{Tr}\left(D(z)'(D(z)^{-1})\right)dz \\ &= N - P.\end{aligned}$$

The last equality follows from the argument principle for the complex function $f(z) := \det(N(z))$ and the Jacobi formula [30] $f'(z)/f(z) = \mathrm{Tr}(N(z)^{-1}N'(z))$.

□

2.4 Analytic system matrix and transfer function

We have seen in (2.1) that a DDS gives a system matrix as well as the transfer function. On the other hand, we now show that a meromorphic matrix can be represented as the transfer function of a system matrix. Our aim in this section is to undertake a comprehensive analysis of system matrices, transfer functions and their canonical forms.

We first discuss transfer function realization of a meromorphic matrix and the system matrix associated with such a realization. The idea is to analyze zeros and poles of a meromorphic matrix via zeros of appropriate holomorphic matrices. For instance, a function $f \in \mathbb{M}(\Omega)$ when represented as $f = p/q$, where $p, q \in \mathbb{H}(\Omega)$ are coprime, can be studied by analyzing the matrix $H(z) := \begin{bmatrix} 0 & -p(z) \\ 1 & q(z)\end{bmatrix}$. Indeed, $f(z)$ is the Schur complement of $q(z)$ in $H(z)$ and $\det H(z) = p(z) = q(z)f(z)$. This shows that

$$\sigma_\Omega(f) = \sigma_\Omega(H) \text{ and } \wp_\Omega(f) = \sigma_\Omega(q).$$

Realization theory for rational matrices is well-known and has been studied extensively [32, 56, 50]. We now define realization of a meromorphic matrix.

Definition 2.4.1. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. Then a representation of* $\mathbf{M}(z)$ *of the form*

$$\mathbf{M}(z) = D(z) + C(z)A(z)^{-1}B(z)$$

is called a realization of $\mathbf{M}(z)$*, where* $(C, A, B) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n}$ *with* $A(z)$ *being regular and* $D \in \mathbb{H}(\Omega)^{m\times n}$*. A realization* (C, A, B, D) *of* $\mathbf{M}(z)$ *is called minimal if* $A(z), C(z)$ *are right coprime and* $A(z), B(z)$ *are left coprime, that is, if*

$$\operatorname{rank} \begin{bmatrix} A(z) \\ C(z) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A(z) & B(z) \end{bmatrix} = r \text{ for all } z \in \Omega.$$

We now associate a holomorphic matrix with a realization of $\mathbf{M}(z)$ which can be gainfully utilized for analyzing zeros and poles of $\mathbf{M}(z)$.

Definition 2.4.2 (system matrix)**.** *Let* $(C, A, B) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n}$ *with* $A(z)$ *being regular and* $D \in \mathbb{H}(\Omega)^{m\times n}$*. A holomorphic matrix* $\mathbf{H} \in \mathbb{H}(\Omega)^{(m+r)\times(n+r)}$ *given by*

$$\mathbf{H}(z) := \left[\begin{array}{c|c} D(z) & C(z) \\ \hline B(z) & -A(z) \end{array}\right]$$

is called a system matrix and $A(z)$ *is called the state matrix. The meromorphic matrix*

$$\mathbf{M}(z) := D(z) + C(z)A(z)^{-1}B(z)$$

is called the transfer function of the system matrix $\mathbf{H}(z)$*. The system matrix* $\mathbf{H}(z)$ *is said to be irreducible if the realization* (C, A, B, D) *is minimal.*

Observe that the transfer function $\mathbf{M}(z)$ is the Schur complement of $-A(z)$ in $\mathbf{H}(z)$. The system matrix $\mathbf{H}(z)$ given in Definition 2.4.2 is called an *analytic matrix description(AMD)* of $\mathbf{M}(z)$. We also refer to the system matrix $\mathbf{H}(z)$ as an analytic system matrix of $\mathbf{M}(z)$.

If $\Omega = \mathbb{C}$ and $\mathbf{M}(z)$ is a rational matrix then the system matrix $\mathbf{H}(z)$ is a matrix polynomial and is referred to as Rosenbrock system matrix [50, 32, 56].

Remark 2.4.3. *We mention that one can define multiple system matrices corresponding to a realization of a meromorphic matrix. Indeed, let* (C, A, B, D) *be a realization of* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. Then*

$$\left[\begin{array}{c|c} D(z) & -C(z) \\ \hline B(z) & A(z) \end{array}\right] \quad \text{and} \quad \left[\begin{array}{c|c} D(z) & C(z) \\ \hline -B(z) & A(z) \end{array}\right]$$

are system matrices and $\mathbf{M}(z) = D(z)+C(z)A(z)^{-1}B(z)$ *is the transfer function. Now, permuting the block rows and columns , we have the system matrices*

$$\left[\begin{array}{c|c} A(z) & -B(z) \\ \hline C(z) & D(z) \end{array}\right], \quad \left[\begin{array}{c|c} A(z) & B(z) \\ \hline -C(z) & D(z) \end{array}\right] \quad \text{and} \quad \left[\begin{array}{c|c} -A(z) & B(z) \\ \hline C(z) & D(z) \end{array}\right]$$

with the same transfer function $\mathbf{M}(z) = D(z) + C(z)A(z)^{-1}B(z)$*. We consider one of the above system matrices as per our convenience.*

The existence of a system matrix of a meromorphic matrix follows from an MFD.

Theorem 2.4.4. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$*. Then there exists an analytic system matrix* $\mathbf{H} \in \mathbb{H}(\Omega)^{(m+r)\times(n+r)}$ *given by*

$$\mathbf{H}(z) := \left[\begin{array}{c|c} D(z) & C(z) \\ \hline B(z) & -A(z) \end{array}\right]$$

for some $r \in \mathbb{N}$ *such that* $\mathbf{M}(z)$ *is the transfer function of* $\mathbf{H}(z)$*.*

Proof. Let $\mathbf{M}(z) = N(z)D(z)^{-1}$ be an MFD of $\mathbf{M}(z)$. Then $\mathbf{H}(z) := \left[\begin{array}{c|c} 0_{m\times n} & N(z) \\ \hline I_n & -D(z) \end{array}\right]$ is an analytic system matrix with transfer function $\mathbf{M}(z) = N(z)D(z)^{-1}$. □

To proceed further, we now define equivalent system matrices.

Definition 2.4.5 (strict system equivalence)**.** *Let* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *be analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} D_1(z) & C_1(z) \\ \hline B_1(z) & -A_1(z) \end{array}\right]_{(m+r)\times(n+r)} \text{and } \mathbf{H}_2(z) := \left[\begin{array}{c|c} D_2(z) & C_2(z) \\ \hline B_2(z) & -A_2(z) \end{array}\right]_{(m+\ell)\times(n+\ell)},$$

where $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$ *are regular. Let* $p \geq \max(r, \ell)$*. Then* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *are said to be strict system equivalent and written as* $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$ *if there exist* $M, N \in \mathrm{GL}_p(\mathbb{H}(\Omega))$ *such that*

$$\left[\begin{array}{c|c} I_m & X(z) \\ \hline 0 & M(z) \end{array}\right] \left[\begin{array}{c|c} \mathbf{H}_1(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \left[\begin{array}{c|c} I_n & 0 \\ \hline Y(z) & N(z) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{H}_2(z) & 0 \\ \hline 0 & I_{p-\ell} \end{array}\right],$$

where $X \in \mathbb{H}(\Omega)^{m\times p}$ *and* $Y \in \mathbb{H}(\Omega)^{p\times n}$*.*

Remark 2.4.6. *We mention that for permuted analytic system matrices, strict system equivalence is defined as follows. Let the system matrices* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *be given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \quad \text{and} \quad \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)}$$

with state matrices $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$. *Let* $p \geq \max(r, \ell)$. *Then* $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$ *if there exist* $M, N \in \mathrm{GL}_p(\mathbb{H}(\Omega))$, $X \in \mathbb{H}(\Omega)^{m\times p}$ *and* $Y \in \mathbb{H}(\Omega)^{p\times n}$ *such that*

$$\left[\begin{array}{c|c} M(z) & 0 \\ \hline X(z) & I_m \end{array}\right] \left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & \mathbf{H}_1(z) \end{array}\right] \left[\begin{array}{c|c} N(z) & Y(z) \\ \hline 0 & I_n \end{array}\right] = \left[\begin{array}{c|c} I_{p-\ell} & 0 \\ \hline 0 & \mathbf{H}_2(z) \end{array}\right].$$

The strict system equivalence defined above reduces to the strict system equivalence introduced by Rosenbrock (see, [50, p.52] and [56, p.58]) when $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are matrix polynomials. In such a case, the polynomial system matrices $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are called Rosenbrock system matrices. We refer to a Rosenbrock system matrix as a polynomial system matrix. It is well known [50, 56] that strict system equivalence of a polynomial system matrix preserves the transfer function. It turns out that the same is true for an analytic system matrix.

Theorem 2.4.7. *Let* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *be analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \quad \text{and} \quad \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)}$$

with state matrices $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$. *If* $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$ *then*

$$\mathbf{M}_1(z) := D_1(z) + C_1(z)A_1(z)^{-1}B_1(z) = D_2(z) + C_2(z)A_2(z)^{-1}B_2(z) =: \mathbf{M}_2(z).$$

Proof. The proof is purely computational and follows from the same arguments as given in the proof of [56, Theorem 2.13] keeping in mind that unimodular matrix polynomials are replaced with invertible holomorphic matrices (unit elements) and matrix polynomials are replaced with holomorphic matrices. □

The strict system equivalence of polynomial system matrices can be characterized by their transfer functions.

Theorem 2.4.8. *[50] Let* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *be irreducible polynomial system matrices. Then* $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$ *if and only if* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *have the same transfer function.*

We show that Theorem 2.4.8 also holds when $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are analytic system matrices. For this purpose, we need to extend Fuhrmann system equivalence [56, p.67] of polynomial system matrices to the case of analytic system matrices. For convenience, we consider a permuted analytic system matrix as given in Remark 2.4.3.

Definition 2.4.9 (Fuhrmann system equivalence). *Let $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ be analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \text{ and } \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)},$$

where $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$ are regular state matrices. Then $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are said to be Fuhrmann system equivalent and written as $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$ if there exist $M \in \mathbb{H}(\Omega)^{\ell\times r}, X \in \mathbb{H}(\Omega)^{m\times r}$ and $N \in \mathbb{H}(\Omega)^{\ell\times r}, Y \in \mathbb{H}(\Omega)^{\ell\times n}$ such that

$$\left[\begin{array}{c|c} M(z) & 0_{\ell\times m} \\ \hline X(z) & I_m \end{array}\right] \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right] = \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right] \left[\begin{array}{c|c} N(z) & Y(z) \\ \hline 0_{n\times r} & I_n \end{array}\right]$$

and the following conditions hold:

(a) $M(z)$ and $A_2(z)$ are left coprime.

(b) $A_1(z)$ and $N(z)$ are right coprime.

Notice that the matrices $M(z)$ and $N(z)$ are non-square when $r \neq \ell$. Even when $r = \ell$, unlike in the case of strict system equivalence, the matrices $M(z)$ and $N(z)$ need not be unit elements of $\mathbb{H}(\Omega)^{r\times r}$. It turns out that, as in the case polynomial system matrices (see,[56, Theorem 2.30]), Fuhrmann system equivalence and strict system equivalence are identical equivalence relations on analytic system matrices.

Theorem 2.4.10. *Let $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ be analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \text{ and } \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)},$$

where $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$ are regular state matrices. Then

$$\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z) \Longleftrightarrow \mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z).$$

Proof. Suppose that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. Then by Definition 2.4.9, we have $M(z)A_1(z) = A_2(z)N(z)$. Further, since A_2 and $M(z)$ are left coprime, by Theorem 2.3.4 there exist $\hat{X} \in \mathbb{H}(\Omega)^{\ell\times\ell}$ and $\hat{Y} \in \mathbb{H}(\Omega)^{r\times\ell}$ such that $A_2(z)\hat{X}(z) + M(z)\hat{Y}(z) = I_\ell$. Similarly, since $A_1(z)$ and $N(z)$ are right coprime, there exist $\tilde{X} \in \mathbb{H}(\Omega)^{r\times\ell}$ and $\tilde{Y} \in \mathbb{H}(\Omega)^{r\times r}$ such that $\tilde{X}(z)N(z) + \tilde{Y}(z)A_1(z) = I_r$. Consequently, we have

$$\begin{bmatrix} -\tilde{X}(z) & \tilde{Y}(z) \\ A_2(z) & M(z) \end{bmatrix} \begin{bmatrix} -N(z) & \hat{X}(z) \\ A_1(z) & \hat{Y}(z) \end{bmatrix} = \begin{bmatrix} I_r & W(z) \\ 0 & I_\ell \end{bmatrix},$$

where $W(z) := -\tilde{X}(z)\hat{X}(z) + \tilde{Y}(z)\hat{Y}(z)$. Since $\begin{bmatrix} I_r & -W(z) \\ 0 & I_\ell \end{bmatrix} \begin{bmatrix} I_r & W(z) \\ 0 & I_\ell \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_\ell \end{bmatrix}$, it follows that $\tilde{X}(z)$ and $\tilde{Y}(z)$ can be chosen such that $W(z) = 0$. Indeed, setting $\tilde{X}(z) \leftarrow -(\tilde{X}(z) + W(z)A_2(z))$ and $\tilde{Y}(z) \leftarrow \tilde{Y}(z) - W(z)M(z)$, we have

$$\begin{bmatrix} -\tilde{X}(z) & \tilde{Y}(z) \\ A_2(z) & M(z) \end{bmatrix} \begin{bmatrix} -N(z) & \hat{X}(z) \\ A_1(z) & \hat{Y}(z) \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_\ell \end{bmatrix} \tag{2.10}$$

which shows that the block matrices in (2.10) are unit elements, that is, the block matrices are holomorphic and invertible.

Now the equality in Definition 2.4.9 can be rewritten as

$$\underbrace{\left[\begin{array}{cc|c} -\tilde{X}(z) & \tilde{Y}(z) & 0 \\ A_2(z) & M(z) & 0 \\ \hline -C_2(z) & X(z) & I_m \end{array}\right]}_{\mathcal{X}(z)} \underbrace{\left[\begin{array}{cc|c} I_\ell & 0 & 0 \\ 0 & A_1(z) & B_1(z) \\ \hline 0 & -C_1(z) & D_1(z) \end{array}\right]}_{I_\ell\oplus\mathbf{H}_1(z)} =$$

$$\underbrace{\left[\begin{array}{cc|c} I_r & 0 & 0 \\ 0 & A_2(z) & B_2(z) \\ \hline 0 & -C_2(z) & D_2(z) \end{array}\right]}_{I_r\oplus\mathbf{H}_2(z)} \underbrace{\left[\begin{array}{cc|c} -\tilde{X}(z) & \tilde{Y}(z)A_1(z) & \tilde{Y}(z)B(z) \\ I_\ell & N(z) & Y(z) \\ \hline 0 & 0 & I_n \end{array}\right]}_{\mathcal{Y}(z)}.$$

Thus, in view of Remark 2.4.6, it follows that $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$ provided that $\mathcal{X}(z)$ and $\mathcal{Y}(z)$ are unit elements. It follows from (2.10) that $\mathcal{X}(z)$ is holomorphic and invertible, that is, $\mathcal{X}(z)$ a unit element. On the other hand,

$$\begin{bmatrix} I_r & \tilde{X}(z) \\ 0 & I_\ell \end{bmatrix} \begin{bmatrix} -\tilde{X}(z) & \tilde{Y}(z)A_1(z) \\ I_r & N(z) \end{bmatrix} = \begin{bmatrix} 0 & I_r \\ I_\ell & N(z) \end{bmatrix} \tag{2.11}$$

shows that the block matrices in (2.11) are holomorphic and invertible which in turn show that $\mathcal{Y}(z)$ is holomorphic and invertible, that is, $\mathcal{Y}(z)$ is a unit element.

Conversely, suppose that $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$. Then by Remark 2.4.6, there exist $M, N \in \mathrm{GL}_p(\mathbb{H}(\Omega))$, $X \in \mathbb{H}(\Omega)^{m\times p}$ and $Y \in \mathbb{H}(\Omega)^{p\times n}$ such that

$$\left[\begin{array}{c|c} M(z) & 0 \\ \hline X(z) & I_m \end{array}\right]\left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & \mathbf{H}_1(z) \end{array}\right] = \left[\begin{array}{c|c} I_{p-\ell} & 0 \\ \hline 0 & \mathbf{H}_2(z) \end{array}\right]\left[\begin{array}{c|c} N(z) & Y(z) \\ \hline 0 & I_n \end{array}\right].$$

Now taking conformal partitions of $M(z), N(z), X(z)$ and $Y(z)$, we have

$$\begin{aligned} &\left[\begin{array}{cc|c} M_{11}(z) & M_{12}(z) & 0 \\ M_{21}(z) & M_{22}(z) & 0 \\ \hline X_1(z) & X_2(z) & I_m \end{array}\right]\left[\begin{array}{cc|c} I_{p-r} & 0 & 0 \\ 0 & A_1(z) & B_1(z) \\ \hline 0 & -C_1(z) & D_1(z) \end{array}\right] = \\ &\left[\begin{array}{cc|c} I_{p-\ell} & 0 & 0 \\ 0 & A_2(z) & B_2(z) \\ \hline 0 & -C_2(z) & D_2(z) \end{array}\right]\left[\begin{array}{cc|c} N_{11}(z) & N_{12}(z) & Y_1(z) \\ N_{21}(z) & N_{22}(z) & Y_2(z) \\ \hline 0 & 0 & I_n \end{array}\right], \end{aligned} \tag{2.12}$$

where $M_{11} \in \mathbb{H}(\Omega)^{(p-\ell)\times(p-r)}$, $M_{22} \in \mathbb{H}(\Omega)^{\ell\times r}$ and $N_{11} \in \mathbb{H}(\Omega)^{(p-\ell)\times(p-r)}$, $N_{22} \in \mathbb{H}(\Omega)^{\ell\times r}$. Now equating the blocks, we have

$$\left[\begin{array}{c|c} M_{22}(z) & 0_{\ell\times m} \\ \hline X_2(z) & I_m \end{array}\right]\left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right] = \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]\left[\begin{array}{c|c} N_{22}(z) & Y_2(z) \\ \hline 0_{n\times r} & I_n \end{array}\right],$$

$M_{12}(z)A_1(z) = N_{12}(z)$ and $M_{21}(z) = A_2(z)N_{21}(z)$. Thus we have

$$\begin{aligned} M(z) &= \begin{bmatrix} M_{11}(z) & M_{12}(z) \\ M_{21}(z) & M_{22}(z) \end{bmatrix} = \begin{bmatrix} M_{11}(z) & M_{12}(z) \\ A_2(z)N_{21}(z) & M_{22}(z) \end{bmatrix}, \\ N(z) &= \begin{bmatrix} N_{11}(z) & N_{12}(z) \\ N_{21}(z) & N_{22}(z) \end{bmatrix} = \begin{bmatrix} N_{11}(z) & M_{12}(z)A_1(z) \\ N_{21}(z) & N_{22}(z) \end{bmatrix}. \end{aligned}$$

Since $M(z)$ and $N(z)$ are unit elements, it follows that $A_2(z)$ and $M_{22}(z)$ are left coprime, and $N_{22}(z)$ and $A_1(z)$ are right coprime, Indeed, let

$$M(z)^{-1} = \begin{bmatrix} X_{11}(z) & X_{12}(z) \\ X_{21}(z) & X_{22}(z) \end{bmatrix} \text{ and } N(z)^{-1} = \begin{bmatrix} Y_{11}(z) & Y_{12}(z) \\ Y_{21}(z) & Y_{22}(z) \end{bmatrix}$$

be conformal partitions. Then

$$\begin{aligned} M(z)M(z)^{-1} &= I_p \Longrightarrow A_2(z)N_{21}(z)X_{12}(z) + M_{22}(z)X_{22}(z) = I_\ell \\ N(z)^{-1}N(z) &= I_p \Longrightarrow Y_{21}(z)M_{12}(z)A_1(z) + Y_{22}(z)N_{22}(z) = I_r. \end{aligned}$$

Hence by Theorem 2.3.4, $A_2(z)$ and $M_{22}(z)$ are left coprime, and $A_1(z)$ and $N_{22}(z)$ are right coprime. This proves that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. □

Let $A, B \in \mathbb{H}(\Omega)^{n\times n}$ be regular. We denote the principal divisor of $\det(A(z))$ by ∂_A and refer to ∂_A as the ***principal divisor*** of $A(z)$ on Ω. Obviously

$$A(z) \sim_\Omega B(z) \Longrightarrow \partial_A = \partial_B.$$

In particular, if $A, B \in \mathbb{C}[z]^{n\times n}$ then the unit elements in $\mathbb{C}[z]^{n\times n}$ are unimodular matrix polynomials and in such a case $A(z) \sim_{\mathbb{C}} B(z)$ is unimodular equivalence. Consequently,

$$\partial_A = \partial_B \Longrightarrow \deg\det(A(z)) = \deg\det(B(z)).$$

Corollary 2.4.11. *Let $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ be analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \text{ and } \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)},$$

where $A_1 \in \mathbb{H}(\Omega)^{r\times r}$ and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$ are regular state matrices. If $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$ then the following hold:

- $\partial_{A_1} = \partial_{A_2}$.
- $\mathbf{M}_1(z) := D_1(z) + C_1(z)A_1(z)^{-1}B_1(z) = D_2(z) + C_2(z)A_2(z)^{-1}B_2(z) =: \mathbf{M}_2(z)$.
- *The non-unit invariant functions in the Smith forms of $A_1(z)$ and $A_2(z)$ are the same.*
- *The non-unit invariant functions in the Smith forms of*

  $$\begin{bmatrix} A_1(z) & B_1(z) \end{bmatrix} \text{ and } \begin{bmatrix} A_2(z) & B_2(z) \end{bmatrix}$$

  are the same.
- *The non-unit invariant functions in the Smith forms of* $\begin{bmatrix} A_1(z) \\ C_1(z) \end{bmatrix}$ *and* $\begin{bmatrix} A_2(z) \\ C_2(z) \end{bmatrix}$ *are the same.*

- *The non-unit invariant functions in the Smith forms of* $\mathbf{H}_1(z)$ *and* $\mathbf{H}_2(z)$ *are the same.*

Proof. By Theorem 2.4.10 we have $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z) \Longleftrightarrow \mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$. Hence equating (1,1) blocks on both sides in (2.12), we have

$$M(z)(I_{p-r}\oplus A_1(z)) = (I_{p-\ell}\oplus A_2(z))N(z) \Longrightarrow \det(M(z))\det(A_1(z)) = \det(N(z))\det(A_2(z)).$$

Since $\det(M(z))$ and $\det(N(z))$ are unit elements of $\mathbb{H}(\Omega)$, we have $\partial_{A_1} = \partial_{A_2}$.

Similarly, the remaining results follow from (2.12) by equating appropriate blocks and their Smith forms. □

Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. Then we have seen that $\mathbf{M}(z)$ admits a right (resp., left) MFD of the form $\mathbf{M}(z) = N_R(z)D_R(z)^{-1}$ (resp., $\mathbf{M}(z) = D_L(z)^{-1}N_L(z)$), where $N_L, N_R \in \mathbb{H}(\Omega)^{m\times n}$ and, $D_L \in \mathbb{H}(\Omega)^{m\times m}$ and $D_R \in \mathbb{H}(\Omega)^{n\times n}$ are regular. It follows that $\mathbf{M}(z)$ is the transfer function of the system matrices

$$\mathbf{H}_R(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix}_{(n+m)\times(n+n)} \text{ and } \mathbf{H}_L(z) := \begin{bmatrix} D_L(z) & N_L(z) \\ -I_m & 0_{m\times n} \end{bmatrix}_{(m+m)\times(m+n)}.$$

A system matrix of the form $\mathbf{H}_R(z)$ (or a block permutation similarity transformation of $\mathbf{H}_R(z)$) is called a system matrix in ***right matrix-fraction form*** which we refer to as ***RMF-system matrix***. Similarly, a system matrix of the form $\mathbf{H}_L(z)$ (or a block permutation similarity transformation of $\mathbf{H}_L(z)$) is called a system matrix in ***left matrix-fraction form*** which we refer to as ***LMF-system matrix***.

Theorem 2.4.12. *Let* $\mathbf{H}_1, \mathbf{H}_2 \in \mathbb{H}(\Omega)^{(n+m)\times(n+n)}$ *be RMF-system matrices given by*

$$\mathbf{H}_1(z) := \begin{bmatrix} D_1(z) & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} \text{ and } \mathbf{H}_2(z) := \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix}.$$

Then $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z) \Longleftrightarrow$ *there exist* $T_R \in \mathrm{GL}_n(\mathbb{H}(\Omega))$ *such that*

$$\begin{bmatrix} D_1(z) & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} T_R(z) & 0 \\ 0 & I_n \end{bmatrix}.$$

Proof. We only need to prove the implication ($\Rightarrow$). Suppose that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. Then there exist $M, N, Y \in \mathbb{H}(\Omega)^{n\times n}$ and $X \in \mathbb{H}(\Omega)^{m\times n}$ such that

$$\begin{bmatrix} M(z) & 0_{n\times m} \\ X(z) & I_m \end{bmatrix} \begin{bmatrix} D_1(z) & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} N(z) & Y(z) \\ 0_{n\times n} & I_n \end{bmatrix} \quad (2.13)$$

and $M(z), D_2(z)$ are left coprime and $D_1(z), N(z)$ are right coprime. Multiplying both sides of (2.13) from the right by $\begin{bmatrix} I_n & 0_{n\times n} \\ -D_1(z) & I_n \end{bmatrix}$ we have

$$\begin{bmatrix} M(z) & 0_{n\times m} \\ X(z) & I_m \end{bmatrix} \begin{bmatrix} 0_{n\times n} & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} T_R(z) & Y(z) \\ -D_1(z) & I_n \end{bmatrix}, \tag{2.14}$$

where $T_R(z) := N(z) - Y(z)D_1(z)$. Now, equating (1, 1) and (2, 1) blocks on both sides of (2.14), we have $D_1(z) = D_2(z)T_R(z)$ and $N_1(z) = N_2(z)T_R(z)$. This shows that

$$\begin{bmatrix} D_1(z) & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} T_R(z) & 0 \\ 0 & I_n \end{bmatrix}.$$

It remains to show that $T_R \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. Since $D_1(z) = D_2(z)T_2(z)$, it follows that $T_R(z)$ is regular and $\partial_{D_1} = \partial_{D_2} + \partial_{T_R}$, where $\partial_{D_1}, \partial_{D_2}$ and ∂_{T_R}, respectively, are the principal divisors of $D_1(z), D_2(z)$ and $T_R(z)$ on Ω. By Corollary 2.4.11, $\partial_{D_1} = \partial_{D_2}$. Hence we have $\partial_{T_R} = 0$. In other words, $\det(T_R(z))$ is a unit element of $\mathbb{H}(\Omega)$ which in turn shows that $T_R \in \mathrm{GL}_n(\mathbb{H}(\Omega))$. □

By similar arguments, we have the following result.

Corollary 2.4.13. *Let* $\mathbf{H}_1, \mathbf{H}_2 \in \mathbb{H}(\Omega)^{(m+m)\times(m+n)}$ *be LMF-system matrices given by*

$$\mathbf{H}_1(z) := \begin{bmatrix} D_1(z) & N_1(z) \\ -I_m & 0_{m\times n} \end{bmatrix} \text{ and } \mathbf{H}_2(z) := \begin{bmatrix} D_2(z) & N_2(z) \\ -I_m & 0_{m\times n} \end{bmatrix}.$$

Then $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z) \iff$ *there exist* $T_L \in \mathrm{GL}_m(\mathbb{H}(\Omega))$ *such that*

$$\begin{bmatrix} D_1(z) & N_1(z) \\ -I_m & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} T_L(z) & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} D_2(z) & N_2(z) \\ -I_m & 0_{m\times n} \end{bmatrix}.$$

We now show that an analytic system matrix, under appropriate assumption, is Fuhrmann system equivalent to an RMF-system matrix.

Theorem 2.4.14. *Let* $(C, A, B) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n}$ *with* $A(z)$ *being regular and* $D \in \mathbb{H}(\Omega)^{m\times n}$. *Consider the system matrix*

$$\mathbf{H}_1(z) := \begin{bmatrix} A(z) & B(z) \\ -C(z) & D(z) \end{bmatrix} \in \mathbb{H}(\Omega)^{(r+m)\times(r+n)}.$$

If $A(z)$ and $B(z)$ are left coprime then there exists an RMF-system matrix

$$\mathbf{H}_2(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix} \in \mathbb{H}(\Omega)^{(n+m)\times(n+n)}$$

such that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. Further, if $A(z)$ and $C(z)$ are right coprime then $N_R(z)$ and $D_R(z)$ are right coprime.

Proof. Since $A(z)$ and $B(z)$ are left coprime, by an analogue of Theorem 2.3.4(e), there exists $N \in \mathbb{H}(\Omega)^{n\times r}$ and $Y \in \mathbb{H}(\Omega)^{n\times n}$ such that

$$T(z) := \begin{bmatrix} A(z) & B(z) \\ -N(z) & -Y(z) \end{bmatrix} \in \mathrm{GL}_{r+n}(\mathbb{H}(\Omega)).$$

Consider the conformal partition of $T(z)^{-1}$ be given by

$$T(z)^{-1} = \begin{bmatrix} T_1(z) & T_2(z) \\ M(z) & D_R(z) \end{bmatrix} \in \mathrm{GL}_{r+n}(\mathbb{H}(\Omega)),$$

where with $T_1 \in \mathbb{H}(\Omega)^{r\times r}$ and $D_R \in \mathbb{H}(\Omega)^{n\times n}$. Then

$$\begin{bmatrix} T_1(z) & T_2(z) \\ M(z) & D_R(z) \end{bmatrix} \begin{bmatrix} A(z) & B(z) \\ -N(z) & -Y(z) \end{bmatrix} = T(z)^{-1}T(z) = \begin{bmatrix} I_r & 0 \\ 0 & I_n \end{bmatrix} \tag{2.15}$$

gives $M(z)A(z) = D_R(z)N(z)$ and $M(z)B(z) = D_R(z)Y(z) + I_n$ which can be written together as

$$M(z)\begin{bmatrix} A(z) & B(z) \end{bmatrix} = \begin{bmatrix} D_R(z) & I_n \end{bmatrix} \begin{bmatrix} N(z) & Y(z) \\ 0_{n\times r} & I_n \end{bmatrix}. \tag{2.16}$$

Again using the fact that $T(z)T(z)^{-1} = I_{r+n}$, we have

$$\begin{bmatrix} A(z) & B(z) \\ -C(z) & D(z) \end{bmatrix} \begin{bmatrix} T_1(z) & T_2(z) \\ M(z) & D_R(z) \end{bmatrix} = \begin{bmatrix} I_r & 0_{r\times n} \\ -X(z) & N_R(z) \end{bmatrix} = \begin{bmatrix} I_r & 0_{r\times m} \\ -X(z) & I_m \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & N_R(z) \end{bmatrix},$$

where $X(z) := C(z)T_1(z) - D(z)M(z)$ and $N_R(z) := D(z)D_R(z) - C(z)T_2(z)$. Hence

$$\begin{aligned} \begin{bmatrix} I_r & 0_{r\times m} \\ X(z) & I_m \end{bmatrix} \begin{bmatrix} A(z) & B(z) \\ -C(z) & D(z) \end{bmatrix} &= \begin{bmatrix} I_r & 0 \\ 0 & N_R(z) \end{bmatrix} \begin{bmatrix} T_1(z) & T_2(z) \\ M(z) & D_R(z) \end{bmatrix}^{-1} \\ &= \begin{bmatrix} I_r & 0 \\ 0 & N_R(z) \end{bmatrix} \begin{bmatrix} A(z) & B(z) \\ -N(z) & -Y(z) \end{bmatrix} \end{aligned}$$

gives (by equating the last block row)

$$X(z)\begin{bmatrix}A(z) & B(z)\end{bmatrix} + \begin{bmatrix}-C(z) & D(z)\end{bmatrix} = -N_R(z)\begin{bmatrix}N(z) & Y(z)\end{bmatrix}. \tag{2.17}$$

Now writing (2.16) and (2.17) together we have

$$\begin{bmatrix}M(z) & 0_{n\times m}\\ X(z) & I_m\end{bmatrix}\begin{bmatrix}A(z) & B(z)\\ -C(z) & D(z)\end{bmatrix} = \begin{bmatrix}D_R(z) & I_n\\ -N_R(z) & 0_{m\times n}\end{bmatrix}\begin{bmatrix}N(z) & Y(z)\\ 0_{n\times r} & I_n\end{bmatrix}.$$

It follows from (2.15) that $A(z)$ and $N(z)$ are right coprime and $M(z)$ and $D_R(z)$ are left coprime. It remains to show that $D_R(z)$ is regular. Let $f \in \mathbb{H}(\Omega)^n$ and $D_R f = 0$. Now $T(z)T(z)^{-1} = I_r \oplus I_n$ gives $AT_2 f + BD_R f = 0 \Longrightarrow AT_2 f = 0 \Longrightarrow T_2 f = 0$ since $A(z)$ is regular. This shows that $\begin{bmatrix}T_2\\ D_R\end{bmatrix} f = 0 \Longrightarrow f = 0$ since $T(z)^{-1}$ is a unit element. Hence $D_R(z)$ is regular. This proves that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$.

By Corollary 2.4.11, the matrices $\begin{bmatrix}A(z)\\ -C(z)\end{bmatrix}$ and $\begin{bmatrix}D_R(z)\\ -N_R(z)\end{bmatrix}$ have the identical non-unit invariant functions in their Smith forms. Thus if $A(z)$ and $C(z)$ are right coprime then $\begin{bmatrix}I_r\\ 0_{m\times r}\end{bmatrix}$ is the Smith form of $\begin{bmatrix}A(z)\\ -C(z)\end{bmatrix}$. Consequently, $\begin{bmatrix}I_n\\ 0_{m\times n}\end{bmatrix}$ is the Smith form of $\begin{bmatrix}D_R(z)\\ -N_R(z)\end{bmatrix}$ which in turn shows that $D_R(z)$ and $N_R(z)$ are right coprime. □

We mention that our proof of Theorem 2.4.14 is concise and simpler than the proof of [56, Theorem 2.40] for polynomial system matrices.

We have the following result whose proof is similar to that of Theorem 2.4.14.

Theorem 2.4.15. *Let $(C, A, B) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n}$ with $A(z)$ being regular and $D \in \mathbb{H}(\Omega)^{m\times n}$. Consider the system matrix*

$$\mathbf{H}_1(z) := \begin{bmatrix}A(z) & B(z)\\ -C(z) & D(z)\end{bmatrix} \in \mathbb{H}(\Omega)^{(r+m)\times(r+n)}.$$

If $A(z)$ and $C(z)$ are right coprime then there exists an LMF-system matrix

$$\mathbf{H}_2(z) := \begin{bmatrix}D_L(z) & N_L(z)\\ -I_m & 0_{m\times n}\end{bmatrix} \in \mathbb{H}(\Omega)^{(m+m)\times(m+n)}$$

such that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$*. Further, if* $A(z)$ *and* $B(z)$ *are left coprime then* $D_L(z)$ *and* $N_L(z)$*are left coprime.*

We are now ready to prove that two irreducible analytic system matrices are strict system equivalent (also Fuhrmann system equivalent) if and only if they have the same transfer function.

Theorem 2.4.16. *Let* $\mathbf{H}_1 \in \mathbb{H}(\Omega)^{(r+m)\times(r+n)}$ *and* $\mathbf{H}_2 \in \mathbb{H}(\Omega)^{(\ell+m)\times\ell+n)}$ *be irreducible analytic system matrices given by*

$$\mathbf{H}_1(z) := \left[\begin{array}{c|c} A_1(z) & B_1(z) \\ \hline -C_1(z) & D_1(z) \end{array}\right]_{(r+m)\times(r+n)} \quad \text{and} \quad \mathbf{H}_2(z) := \left[\begin{array}{c|c} A_2(z) & B_2(z) \\ \hline -C_2(z) & D_2(z) \end{array}\right]_{(\ell+m)\times(\ell+n)}$$

with transfer functions $\mathbf{M}_i(z) := D_i(z) + C_i(z)A_i(z)^{-1}B_i(z),\ i = 1, 2,$ *where* $A_1 \in \mathbb{H}(\Omega)^{r\times r}$and $A_2 \in \mathbb{H}(\Omega)^{\ell\times\ell}$ *are regular state matrices. Then*

$$\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z) \Longleftrightarrow \mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z) \Longleftrightarrow \mathbf{M}_1(z) = \mathbf{M}_2(z).$$

Proof. We have already seen that $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z) \Longleftrightarrow \mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$ and in such a case, by Theorem 2.4.7 and Corollary 2.4.11, we have $\mathbf{M}_1(z) = \mathbf{M}_2(z)$.

Conversely, suppose that $\mathbf{M}_1(z) = \mathbf{M}_2(z)$. By Theorem 2.4.14, there exist RMF-system matrices

$$\mathbf{S}_i(z) := \begin{bmatrix} D_i(z) & I_n \\ -N_i(z) & 0_{m\times n} \end{bmatrix} \in \mathbb{H}(\Omega)^{(n+m)\times(n+n)}$$

with transfer function $N_i(z)D_i(z)^{-1}$ such that $\mathbf{H}_i(z) \sim_{fse} \mathbf{S}_i(z)$ for $i = 1, 2$. Hence by Corollary 2.4.11, $\mathbf{M}_i(z) = N_i(z)D_i(z)^{-1}$ for $i = 1, 2$. Now

$$\mathbf{M}_1(z) = \mathbf{M}_2(z) \Longrightarrow N_1(z)D_1(z)^{-1} = N_2(z)D_2(z)^{-1}.$$

Since $\mathbf{H}_i(z)$ is irreducible, by Theorem 2.4.14 the MFD $N_i(z)D_i(z)^{-1}$ is right coprime for $i = 1, 2$. Hence by Theorem 2.3.8, there exist $T_R \in \mathrm{GL}_n(\mathbb{H}(\Omega))$ such that

$$N_1(z) = N_2(z)T_R(z) \text{ and } D_1(z) = D_2(z)T_R(z).$$

This shows that

$$\begin{bmatrix} D_1(z) & I_n \\ -N_1(z) & 0_{m\times n} \end{bmatrix} = \begin{bmatrix} D_2(z) & I_n \\ -N_2(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} T_R(z) & 0 \\ 0 & I_n \end{bmatrix}.$$

Now, by Theorem 2.4.12, we have $\mathbf{S}_1(z) \sim_{fse} \mathbf{S}_2(z)$. Consequently, we have

$$\mathbf{H}_1(z) \sim_{fse} \mathbf{S}_1(z) \sim_{fse} \mathbf{S}_2(z) \sim_{fse} \mathbf{H}_2(z),$$

that is, $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. This completes the proof. □

We now examine the conditions under which an LMF-system matrix is Fuhrmann system equivalent to an RMF-system matrix.

Theorem 2.4.17. *Consider the LMF-system matrix $\mathbf{H}_L(z) \in \mathbb{H}(\Omega)^{(m+m)\times(m+n)}$ and the RMF-system matrix $\mathbf{H}_R(z) \in \mathbb{H}(\Omega)^{(n+m)\times(n+n)}$ given by*

$$\mathbf{H}_L(z) := \begin{bmatrix} D_L(z) & N_L \\ -I_m & 0_{m\times n} \end{bmatrix} \text{ and } \mathbf{H}_R(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix}.$$

Then

$$\mathbf{H}_L(z) \sim_{fse} \mathbf{H}_R(z) \Longleftrightarrow \begin{cases} (a)\ D_L(z)^{-1}N_L(z) = N_R(z)D_R(z)^{-1} \\ (b)\ D_L(z), N_L(z) \text{ are left coprime} \\ (c)\ D_R(z), N_R(z) \text{ are right coprime.} \end{cases}$$

Proof. First, note that $D_L(z)^{-1}N_L(z)$ is the transfer function of $\mathbf{H}_L(z)$ and $N_R(z)D_L(z)^{-1}$ is the transfer function of $\mathbf{H}_R(z)$. Thus, if the conditions (a), (b) and (c) hold then $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are irreducible and have the same transfer function. Hence by Theorem 2.4.16, we have $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$.

Conversely, suppose that $\mathbf{H}_1(z) \sim_{fse} \mathbf{H}_2(z)$. Then by Corollary 2.4.11, the non-unit invariant functions in the Smith forms of $\begin{bmatrix} D_L(z) & N_L(z) \end{bmatrix}$ and $\begin{bmatrix} D_R(z) & I_n \end{bmatrix}$ are identical. Also, the non-unit invariant functions in the Smith forms of $\begin{bmatrix} D_L \\ -I_m \end{bmatrix}$ and $\begin{bmatrix} D_R \\ -N_R(z) \end{bmatrix}$ are identical. Since $D_R(z)$ and I_n are left coprime, the invariant functions in the Smith form of $\begin{bmatrix} D_R(z) & I_n \end{bmatrix}$ are all unit elements. This implies that the invariant functions in the Smith form of $\begin{bmatrix} D_L(z) & N_L(z) \end{bmatrix}$ are all unit elements. Hence $D_L(z)$ and $N_L(z)$ are left coprime. Similarly $D_L(z)$ and I_m are right coprime which implies that $D_R(z)$ and $N_R(z)$ are right coprime. Again, by Corollary 2.4.11, we have $D_L^{-1}N_L(z) = N_R(z)D_R(z)^{-1}$. □

We now examine relationships between canonical forms of analytic system matrix, state matrix and the transfer function.

Theorem 2.4.18. *Let* $(C, A, B) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n}$ *with* $A(z)$ *being regular and* $D \in \mathbb{H}(\Omega)^{m\times n}$*. Consider the analytic system matrix*

$$\mathbf{H}(z) := \begin{bmatrix} A(z) & B(z) \\ -C(z) & D(z) \end{bmatrix} \in \mathbb{H}(\Omega)^{(r+m)\times(r+n)}$$

and its transfer function $\mathbf{M}(z) := D(z) + C(z)A(z)^{-1}B(z) \in \mathbb{H}(\Omega)^{m\times n}$*. Suppose that* $\mathbf{H}(z)$ *is irreducible and* $\mathrm{nrank}(\mathbf{M}) = p$*. Let*

$$\Sigma_{\mathbf{M}}(z) := \phi_1(z)/\psi_1(z) \oplus \cdots \oplus \phi_p(z)/\psi_p(z) \oplus 0_{m-p,n-p}$$

be the Smith-McMillan form of $\mathbf{M}(z)$*. Then*

$$\begin{aligned} S_A(z) &:= I_{r-p} \oplus \mathrm{diag}(\psi_p(z), \psi_{p-1}(z), \psi_1(z)) \\ S_{\mathbf{H}}(z) &:= I_r \oplus \mathrm{diag}(\phi_1(z), \dots, \phi_p(z)) \oplus 0_{m-p,n-p} \end{aligned}$$

are the Smith forms of $A(z)$ *and* $\mathbf{H}(z)$*, respectively. In particular, we have*

- $\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{H})$ *and* $\mathrm{Ind}_e(\lambda, \mathbf{M}) = \mathrm{Ind}_e(\lambda, \mathbf{H})$ *for all* $\lambda \in \sigma_\Omega(\mathbf{M})$*.*
- $\wp_\Omega(\mathbf{M}) = \sigma_\Omega(A)$ *and* $\mathrm{Ind}_p(\lambda, \mathbf{M}) = \mathrm{Ind}_e(\lambda, A)$ *for all* $\lambda \in \wp_\Omega(\mathbf{M})$*.*

Proof. Since $\mathbf{H}(z)$ is irreducible, by Theorem 2.4.14 there exists an irreducible RMF-system matrix (with $D_R(z), N_R(z)$ are right coprime)

$$\mathbf{S}(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix} \in \mathbb{H}(\Omega)^{(n+m)\times(n+n)}$$

such that $\mathbf{H}(z) \sim_{fse} \mathbf{S}(z)$. By Corollary 2.4.11, $\mathbf{M}(z) = N_R(z)D_R(z)^{-1}$ is a right coprime MFD. Now by Theorem 2.3.9, $S_{D_R}(z) := I_{n-p} \oplus \mathrm{diag}(\psi_p(z), \psi_{p-1}(z), \psi_1(z))$ is the Smith form of $D_R(z)$ and $S_{N_R}(z) := \mathrm{diag}(\phi_1(z), \dots, \phi_p(z)) \oplus 0_{m-p,n-p}$ is the Smith form of $N_R(z)$.

By Corollary 2.4.11, the non-unit invariant functions in the Smith forms of $A(z)$ and $D_R(z)$ are the same. Hence $S_A(z) = I_{r-p} \oplus \mathrm{diag}(\psi_p(z), \psi_{p-1}(z), \psi_1(z))$ is the Smith form of $A(z)$. Next, we have

$$\begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix} \begin{bmatrix} 0_{n\times n} & -I_n \\ I_n & D_R(z) \end{bmatrix} = \begin{bmatrix} I_n & 0_{n\times n} \\ 0_{m\times n} & N_R(z) \end{bmatrix} =: \mathbf{H}_1(z),$$

where the second matrix on the left hand side is a unit element. This shows that $\mathbf{S}(z)$ and $\mathbf{H}_1(z)$ have the same Smith form and $S_{\mathbf{H}_1}(z) = I_n \oplus \mathrm{diag}(\phi_1(z), \dots, \phi_p(z)) \oplus 0_{m-p,n-p}$

is the Smith form of $\mathbf{H}_1(z)$. By Corollary 2.4.11, the non-unit invariant functions in the Smith forms of $\mathbf{H}(z)$ and $\mathbf{S}(z)$ are the same. Consequently,

$$S_{\mathbf{H}}(z) = I_r \oplus \operatorname{diag}(\phi_1(z), \ldots, \phi_p(z)) \oplus 0_{m-p,n-p}$$

is the Smith form of $\mathbf{H}(z)$. The proofs of the remaining results are immediate. $\square$

Finally, we have the following result that establishes equivalence between an irreducible system matrix and an irreducible RFM system matrix. Obviously, similar result holds for an irreducible LMF-system matrix.

Theorem 2.4.19. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. Let $\mathbf{M}(z) = N_R(z)D_R(z)^{-1}$ be a right coprime MFD, where $D_R(z)$ is regular and $(N_R, D_R) \in \mathbb{H}(\Omega)^{m\times n} \times \mathbb{H}(\Omega)^{n\times n}$. Let*

$$\mathbf{M}(z) = D(z) + C(z)A(z)^{-1}B(z)$$

be a minimal realization, where $(C, A, B, D) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times r} \times \mathbb{H}(\Omega)^{r\times n} \times \mathbb{H}(\Omega)^{m\times n}$ and $A(z)$ is regular. Consider the system matrix $\mathbf{H}(z)$ and the RMF-system matrix $\mathbf{S}(z)$ given by

$$\mathbf{H}(z) := \begin{bmatrix} A(z) & B(z) \\ -C(z) & D(z) \end{bmatrix} \text{ and } \mathbf{S}(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix}.$$

Let $p \geq \max(r, n)$. Then the following hold:

- $\mathbf{H}(z) \sim_{fse} \mathbf{S}(z)$ *or equivalently* $\mathbf{H}(z) \sim_{sse} \mathbf{S}(z)$.
- $I_{p-r} \oplus \mathbf{H}(z) \sim_{sse} I_{p-n} \oplus \mathbf{S}(z) \sim_{\Omega} I_p \oplus N_R(z)$.
- $I_{p-r} \oplus A(z) \sim_{\Omega} I_{p-n} \oplus D_R(z)$.

Let $S_{\mathbf{H}}(z), S_{\mathbf{S}}(z)$ and $S_{N_R}(z)$ be the Smith forms of $\mathbf{H}(z), \mathbf{S}(z)$ and $\mathbf{N}_R(z)$, respectively. Then $I_{p-r} \oplus S_A(z) = I_{p-n} \oplus S_{D_R}(z)$ and $I_{p-r} \oplus S_{\mathbf{H}}(z) = I_{p-n} \oplus S_{\mathbf{S}}(z) = I_p \oplus S_{N_R}(z)$.

In particular, $\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{H}) = \sigma_\Omega(\mathbf{S}) = \sigma_\Omega(N_R)$ and $\wp_\Omega(\mathbf{M}) = \sigma_\Omega(A) = \sigma_\Omega(D_R)$. Further, $\operatorname{ind}_e(\lambda, \mathbf{M}) = \operatorname{ind}_e(\lambda, \mathbf{H}) = \operatorname{ind}_e(\lambda, \mathbf{S}) = \operatorname{ind}_e(\lambda, N_R)$ for $\lambda \in \sigma_\Omega(\mathbf{M})$ and $\operatorname{ind}_p(\lambda, \mathbf{M}) = \operatorname{ind}_e(\lambda, A) = \operatorname{ind}_e(\lambda, D_R)$ for $\lambda \in \wp_\Omega(\mathbf{M})$.

Proof. Since $\mathbf{H}(z)$ and $\mathbf{S}(z)$ are irreducible analytic system matrices having the same transfer function, by Theorem 2.4.16, we have $\mathbf{H}(z) \sim_{fse} \mathbf{S}(z)$. By Theorem 2.4.10, we have $\mathbf{H}(z) \sim_{sse} \mathbf{S}(z)$. Since $\mathbf{H}_1(z) \sim_{sse} \mathbf{H}_2(z)$, by Remark 2.4.6, there exist $M, N \in \mathrm{GL}_p(\mathbb{H}(\Omega))$, $X \in \mathbb{H}(\Omega)^{m\times p}$ and $Y \in \mathbb{H}(\Omega)^{p\times n}$ such that

$$\left[\begin{array}{c|c} M(z) & 0 \\ \hline X(z) & I_m \end{array}\right] \left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & \mathbf{H}(z) \end{array}\right] \left[\begin{array}{c|c} N(z) & Y(z) \\ \hline 0 & I_n \end{array}\right] = \left[\begin{array}{c|c} I_{p-n} & 0 \\ \hline 0 & \mathbf{S}(z) \end{array}\right]. \tag{2.18}$$

This shows that $I_{p-r} \oplus \mathbf{H}(z) \sim_{sse} I_{p-n} \oplus \mathbf{S}(z)$. Now elementary block column operations show that

$$\mathbf{S}(z) := \begin{bmatrix} D_R(z) & I_n \\ -N_R(z) & 0_{m\times n} \end{bmatrix} \sim_\Omega \begin{bmatrix} I_n & 0 \\ 0 & N_R(z) \end{bmatrix} \Longrightarrow I_{p-n} \oplus \mathbf{S}(z) \sim_\Omega I_p \oplus N_R(z).$$

Now equating (1, 1) block on both sides of (2.18), we have

$$M(z)(I_{p-r} \oplus A(z))N(z) = I_{p-n} \oplus D_R(z) \Longrightarrow I_{p-r} \oplus A(z) \sim_\Omega I_{p-n} \oplus D_R(z).$$

The proofs of remaining assertions follow immediately from Theorem 2.3.9 and Theorem 2.4.18. □

3

Spectral perturbation theory for holomorphic operator-valued functions

Let X be a complex Banach space and $L(X)$ be the Banach space of all bounded linear operators on X. Let $\Omega \subset \mathbb{C}$ be open and connected. Let $T, V : \Omega \longrightarrow L(X)$ be holomorphic operator-valued functions. We consider the one parameter family of operator-valued functions $W(\lambda, t) := T(\lambda) + tV(\lambda)$, for $t \in \mathbb{C}$, and analyze evolution of the discrete eigenvalues of $W(\lambda, t)$ when t varies in $\mathbb{C}$. We provide a brief review of the discrete spectrum of $T(\lambda)$ and present several equivalent characterizations for discrete eigenvalues of $T(\lambda)$. We also prove Rouche's theorem for operator-valued functions under a weaker assumption, which we utilize to derive perturbation bounds for the discrete eigenvalues of $W(\lambda, t)$ when $|t|$ is small.

3.1 Introduction

Let X be a complex Banach space and $L(X)$ be the Banach space of all bounded linear operators on X. Let $\Omega \subset \mathbb{C}$ be open and connected. Let $T : \Omega \longrightarrow L(X)$ be holomorphic and regular, that is, $T(\lambda)$ is invertible for some $\lambda \in \Omega$. The nonlinear eigenvalue problem is to solve the equation

$$T(\mu)v = 0$$

for $\mu \in \Omega$ and a nonzero vector $v \in X$. Nonlinear eigenvalue problems arise in many applications and the numerics of nonlinear eigenvalue problems is an active area of research; see [10, 11, 42, 43, 44, 29, 45] and the references therein. For example, the nonlinear eigenvalue problem

$$T(\lambda)v := (\lambda I - A_0 - \sum_{i=1}^{m} A_i e^{-\lambda \tau_i})v = 0$$

arises in the study of the delay differential equation

$$\frac{dx}{dt}(t) = A_0 x(t) + \sum_{i=1}^{m} A_i x(t - \tau_i),$$

where $x(t) \in \mathbb{R}^n$ is the state variable at time t, A_i's are $n \times n$ matrices, and $0 < \tau_1 < \tau_2 < \cdots < \tau_m$ represent the time-delays [45]. The spectral analysis of operator-valued functions such as $T(\lambda)$ has been studied extensively; see [23, 27, 28, 44] and the references therein. Since a nonlinear eigenvalue problem is difficult to solve, an approximation of the nonlinear eigenvalue problem by a simpler one that is easier to solve is a favoured option. Thus, often the operator-valued function $T(\lambda)$ is approximated or perturbed by another operator-valued function $V : \Omega \longrightarrow L(X)$ so that the perturbed nonlinear eigenvalue problem

$$(T(\mu) + V(\mu))u = 0$$

is easier to solve. Spectral perturbation theory for linear operators is a classical subject and has been studied extensively over the years; see for instance [39, 33]. However, spectral perturbation theory for operator-valued functions has not received much attention. The purpose of this study is to fill this gap.

The main aim of this chapter is to analyze *discrete spectra* of the one parameter family of holomorphic operator-valued function $W(\lambda, t) := T(\lambda) + tV(\lambda)$ for $t \in \mathbb{C}$. First, we present a brief review of the discrete spectrum of $T(\lambda)$ and discuss several equivalent characterizations of the discrete spectrum of $T(\lambda)$. Second, Rouche's theorem states that if $f, g : \Omega \longrightarrow \mathbb{C}$ are holomorphic and Γ is a simple closed rectifiable curve and if $\max_{z\in\Gamma} |g(z)f(z)^{-1}| < 1$ then f and $f + g$ have the same number (counting multiplicities) of zeros inside the curve Γ. An operator analogue of Rouche's theorem is proved in [28, 23] which states that if $T(\lambda)$ and $V(\lambda)$ are Fredholm for all $\lambda \in \Omega$ and if $\max_{z\in\Gamma} \|V(z)T(z)^{-1}\| < 1$ then $T(\lambda) + V(\lambda)$ is Fredholm for all $\lambda \in \mathrm{Int}(\Gamma)$, the region enclosed by the curve Γ, and that $T(\lambda)$ and $T(\lambda) + V(\lambda)$ have the same number of eigenvalues (counting multiplicities) in $\mathrm{Int}(\Gamma)$. We show that for discrete eigenvalues, Rouche's theorem holds under a weaker assumption. In fact, for bounded operator-valued functions $T(\lambda)$ and $V(\lambda)$, we prove the following: Suppose that Γ encloses only discrete eigenvalues of $T(\lambda)$ and that $\max_{z\in\Gamma} r_\sigma(V(z)T(z)^{-1}) < 1$, where $r_\sigma(A)$ is the spectral radius of an operator A. Then both $T(\lambda)$ and $T(\lambda) + V(\lambda)$ are Fredholm operators of index zero for all $\lambda \in \mathrm{Int}(\Gamma)$ and that $T(\lambda)$ and $T(\lambda) + V(\lambda)$ have the same number (counting multiplicities) of (discrete) eigenvalues in $\mathrm{Int}(\Gamma)$. Third, if μ is a discrete eigenvalue of $T(\lambda)$ of algebraic multiplicity ℓ enclosed by Γ then we determine a disc $\partial_\Gamma \subset \mathbb{C}$ and show that $\mathrm{Int}(\Gamma)$ contains exactly ℓ discrete eigenvalues (counting multiplicities) $\mu_1(t), \ldots, \mu_\ell(t)$ of $W(\lambda, t)$ for all $t \in \partial_\Gamma$. Further, we show that the map $t \longmapsto \mu_{\text{av}}(t) := (\mu_1(t) + \cdots + \mu_\ell(t))/\ell$ is holomorphic on ∂_Γ. Finally, if ν is the ascent of μ then we show that there exist positive real numbers δ, τ and real constants α, β (independent of t and V) such that

$$\begin{aligned} \|tV\| < \delta \quad &\Longrightarrow \quad |\mu_j(t) - \mu|^\nu \leq \alpha \|tV\| \ \text{ for } \ j = 1, 2, \ldots, \ell, \\ \|tV\| < \tau \quad &\Longrightarrow \quad |\mu_{\text{av}}(t) - \mu| \leq \beta \|tV\|. \end{aligned}$$

We also obtain estimates of τ and β when μ is a simple eigenvalue of $T(\lambda)$.

3.2 Discrete spectra of bounded operators

Let X and Y be complex Banach spaces and let $L(X, Y)$ be the Banach space of all bounded linear transformations from X to Y. We denote $L(X, X)$ by $L(X)$. Let $T \in L(X)$. Let $\rho(T)$ and $\sigma(T)$ denote the resolvent set and spectrum of T given by

$$\rho(T) := \{\lambda \in \mathbb{C} \ : \ T - \lambda I \text{ is invertible}\} \text{ and } \sigma(T) := \mathbb{C} \setminus \rho(T).$$

Let $\Gamma \subset \rho(T)$ be a simple closed positively oriented rectifiable curve. We denote by $\text{Int}(\Gamma)$ the inner domain enclosed by Γ. Let $\lambda \in \text{Int}(\Gamma)$. Define (see [39, 33])

$$P := \frac{1}{2\pi i} \int_\Gamma (zI - T)^{-1} dz \quad \text{and} \quad S := \frac{1}{2\pi i} \int_\Gamma \frac{(T - zI)^{-1}}{z - \lambda} dz.$$

Then P is called the *spectral projection* associated with T and Γ. If $\text{Int}(\Gamma) \cap \sigma(T) = \{\lambda\}$ then we denote P by P_λ and refer to P_λ as the spectral projection associated with T and λ. More generally, if $\sigma := \text{Int}(\Gamma) \cap \sigma(T)$ then P_σ is the spectral projection associated with T and σ. The operator S is called the *reduced resolvent* of $T - \lambda I$.

If $\text{Int}(\Gamma) \cap \sigma(T) = \{\lambda\}$ then in a neighborhood of λ the following Laurant series expansion holds [39, 33]

$$(T - zI)^{-1} = \sum_{j=0}^{\infty} S^{j+1} (z - \lambda)^j - \frac{P}{z - \lambda} - \sum_{j=1}^{\infty} \frac{D^j}{(z - \lambda)^{j+1}}, \tag{3.1}$$

where $D := (T - \lambda I)P$.

Definition 3.2.1. *[39, p.100] Let $\lambda \in \sigma(T)$. Then λ is said to be a discrete eigenvalue of T if λ is an isolated point of $\sigma(T)$ and the spectral projection P_λ associated with T and λ has finite rank. In such a case, $m(\lambda, T) := \text{rank}(P_\lambda)$ is called the algebraic multiplicity of λ. The set $\sigma_d(T)$ of all discrete eigenvalues of T is called the discrete spectrum of T. Let $\lambda \in \sigma_d(T)$. If λ is a pole of order ν of $(T - zI)^{-1}$ then ν is called the ascent of λ.*

It follows from (4.1) that ν is the ascent of λ if and only if $D^{\nu-1} \neq 0$ and $D^{\nu} = 0$. Note that if $\sigma := \mathrm{Int}(\Gamma) \cap \sigma(T)$ and $m(\sigma, T) := \mathrm{rank}(P_\sigma)$ is finite then σ consists of at most $m(\sigma, T)$ discrete eigenvalues of T of total algebraic multiplicity $m(\sigma, T)$. With a view to defining discrete spectrum of operator-valued functions, we now discuss characterizations of discrete spectrum. We proceed as follows.

Let X and Y be Banach spaces and $T \in L(X, Y)$. We denote by $N(T)$ and $R(T)$ the kernel and the range space of T, respectively, that is, $N(T) := \{x \in X : Tx = 0\}$ and $R(T) := \{Tx : x \in X\}$. The operator T is said to be *Fredholm* if $R(T)$ is closed and the quotient space $Y/R(T)$ and $N(T)$ are finite dimensional. The index of a Fredholm operator T is defined by $\mathrm{ind}(T) := \dim(N(T)) - \dim(Y/R(T))$; see [23, 33].

Definition 3.2.2. *[23, 28] Let $U \subset \mathbb{C}$ be open and $W : U \longrightarrow L(X)$ be meromorphic. Let $\lambda \in U$. Then $W(z)$ is said to be finitely meromorphic at λ if λ is a pole of $W(z)$ and there is a $\delta > 0$ such that $0 < |z - \lambda| < \delta \Longrightarrow W(z) = \sum_{j=-\ell}^{\infty} (z - \lambda)^j A_j$, where the operators $A_j, j = -1, -2, \ldots, -\ell$, are of finite rank.*

Theorem 3.2.3. *Let $\lambda \in \sigma(T)$ be an isolated point. Then the following conditions are equivalent.*

(a) λ is a discrete eigenvalue of T.
(b) $T - \lambda I$ is Fredholm and $\mathrm{ind}(T - \lambda I) = 0$.
(c) $T - \lambda I$ is Fredholm.
(d) There is a $\delta > 0$ such that $|z - \lambda| < \delta \Longrightarrow T - zI$ is Fredholm.
(e) $(T - zI)^{-1}$ is finitely meromorphic at λ.

Proof. Suppose that (a) holds. Then by [33, Theorem IV.5.28], $T - \lambda I$ is Fredholm. Since λ is an isolated point, there is a neighbourhood $\mathrm{Nbd}(\lambda)$ of λ such that $T - zI$ is invertible on $\mathrm{Nbd}(\lambda) \setminus \{\lambda\}$. Since the Fredholm index is continuous [33], it follows that $\mathrm{ind}(T - \lambda I) = 0$. This proves (b).

Now (b) $\Rightarrow$ (c) is immediate. Next, suppose that (c) holds. Since the set of Fredholm operators is an open set in $L(X)$, it follows that there is a neighbourhood $\mathrm{Nbd}(\lambda)$ of λ such that $T - zI$ is Fredholm for each $z \in \mathrm{Nbd}(\lambda)$. This proves (d).

Now suppose that (d) holds. Then, as we shall see in Theorem 3.3.5(a), $(T - zI)^{-1}$ is finitely meromorphic. This proves (e).

Finally, suppose that (e) holds. Then for z close to λ, we have $(T - zI)^{-1} = \sum_{j=-\ell}^{\infty} (z - \lambda)^j A_j$, where the operators $A_j, j = -1, -2, \ldots, -\ell$, are of finite rank. Comparing with the Laurent series (4.1), it follows that $P = -A_{-1}$. This shows that $\mathrm{rank}(P) = \mathrm{rank}(A_{-1}) < \infty$. Hence λ is a discrete eigenvalue of T. This proves (a). $\square$

Remark 3.2.4. *If $(T - zI)^{-1}$ is finitely meromorphic at λ then by (4.1), we have $(T - zI)^{-1} = \sum_{j=0}^{\infty} S^{j+1} (z - \lambda)^j - \frac{P}{z - \lambda} - \sum_{j=1}^{\ell-1} D^j (z - \lambda)^{-(j+1)}$. It is well known [39,* *Proposition 6.2] that $SP = 0$ and $S(T - \lambda I) = (T - \lambda I)S = I - P$. Since $\operatorname{rank}(P) < \infty$, it follows that S is Fredholm and $\operatorname{ind}(S) = 0$. This shows that if $(T - zI)^{-1}$ is finitely meromorphic at λ then the constant term in the Laurent series expansion of $(T - zI)^{-1}$ at λ is a Fredholm operator of index zero.*

3.3 Discrete spectra of operator-valued functions

Let $U \subset \mathbb{C}$ be open. We consider a regular holomorphic operator-valued function $T : U \longrightarrow L(X)$ and analyze its discrete spectrum. The spectrum and the discrete spectrum of $T(\lambda)$ are defined as follows.

Definition 3.3.1. *Let $T : U \longrightarrow L(X)$ be holomorphic. Then the operator-valued function $T(\lambda)$ is said to be regular if the operator $T(\mu)$ is invertible for some $\mu \in U$. Suppose that $T(\lambda)$ is regular.*

(a)The resolvent set $\rho(T)$ and spectrum $\sigma(T)$ of the operator-valued function $T(\lambda)$ are defined by

$$\rho(T) := \{\mu \in U : T(\mu) \ \text{ is invertible}\} \text{ and } \sigma(T) := U \setminus \rho(T).$$

(b) Let $\mu \in \sigma(T)$. Then μ is said to be an eigenvalue of $T(\lambda)$ if there is a nonzero vector $v \in X$ such that $T(\mu)v = 0$. The vector v is called an eigenvector of $T(\lambda)$ corresponding to the eigenvalue μ.

(c) Let $\mu \in \sigma(T)$ be an eigenvalue. Then μ is said to be a discrete eigenvalue of $T(\lambda)$ if μ is an isolated point of $\sigma(T)$ and $T(\mu)$ is Fredholm. The discrete spectrum $\sigma_d(T)$ of $T(\lambda)$ is the set of all discrete eigenvalues of $T(\lambda)$.

(d) Let $\mu \in \sigma_d(T)$. If μ is a pole of $T(\lambda)^{-1}$ order ν then ν is called the ascent of the eigenvalue μ.

Note that $T(\lambda)$ is regular $\iff$ $\rho(T) \neq \emptyset$. We always assume that $T(\lambda)$ is regular and $\Gamma \subset \rho(T)$ is a simple closed positively oriented rectifiable curve.

The concept of equivalence and extension, which will allow us to compare spectra of different operator-valued functions, are defined as follows. Let X_i and Y_i be Banach spaces, $i = 1, 2$.

Definition 3.3.2. *[23, p.38] Let $S : U \to L(X_1, Y_1)$ and $T : U \to L(X_2, Y_2)$ be operator valued functions. Then $S(\lambda)$ and $T(\lambda)$ are called (globally) equivalent on U if there exist operator valued functions $E : U \to L(X_1, X_2)$ and $F : U \to L(Y_2, Y_1)$, which are analytic on U, such that*

$$S(\lambda) = F(\lambda)T(\lambda)E(\lambda), \quad \lambda \in U,$$

and, in addition, $E(\lambda)$ and $F(\lambda)$ are invertible for each $\lambda \in \Omega$. In that case also

$$T(\lambda) = F(\lambda)^{-1} S(\lambda) E(\lambda)^{-1}, \ \lambda \in U,$$

and the operator-valued functions $E(\lambda)^{-1}$ and $F(\lambda)^{-1}$ are again analytic on U.

We write $S(\lambda) \sim_U T(\lambda)$ when $S(\lambda)$ and $T(\lambda)$ are equivalent on U. On the other hand, $S(\lambda)$ and $T(\lambda)$ are said to be (locally) equivalent at $\mu \in U$ if there is a neighbourhood $\mathrm{Nbd}(\mu)$ of μ such that $S(\lambda)$ and $T(\lambda)$ are equivalent on $\mathrm{Nbd}(\mu)$; see [23, p.42]. We write $S(\lambda) \sim_\mu T(\lambda)$ when $S(\lambda)$ and $T(\lambda)$ are equivalent at μ.

Definition 3.3.3. *[23, p.38] Let X and Y be Banach spaces. Given an operator-valued function $A : U \to L(X, Y)$ and a Banach space Z, the operator-valued function*

$$U \longrightarrow L(X \oplus Z, Y \oplus Z), \quad \lambda \longmapsto \begin{bmatrix} A(\lambda) & 0 \\ 0 & I_Z \end{bmatrix},$$

is called the Z-extension of $A(\lambda)$, where I_Z is the identity operator on Z. We write the Z-extension of $A(\lambda)$ by $A(\lambda) \oplus I_Z$.

Observe that if $S : U \longrightarrow L(X)$ is holomorphic and $T(\lambda) \sim_U S(\lambda)$ then $S(\lambda)$ is regular and $\sigma(T) = \sigma(S)$ when $T(\lambda)$ is regular. Thus the regularity of $T(\lambda)$ and the spectrum $\sigma(T)$ do not change when $T(\lambda)$ is replaced by an equivalent operator-valued function. The same conclusions hold when $T(\lambda)$ is replaced with the Z-extension $T(\lambda) \oplus I_Z$. We consider only regular holomorphic operator-valued functions.

We now collect some basic results which will be used later.

Lemma 3.3.4. *Let $U \subset \mathbb{C}$ be open and connected. Let $S_1, S_2 : U \longrightarrow L(X)$ be holomorphic and regular. Let $\mu \in U$. Also, let Z be a Banach space. Then we have the following.*

(a) Suppose that $S_1(\lambda) \sim_\mu S_2(\lambda)$. If $S_1(\mu)$ is Fredholm then $S_2(\mu)$ is Fredholm and $\mathrm{ind}(S_1(\mu)) = \mathrm{ind}(S_2(\mu))$. Also $S_1(\mu) \oplus I_Z$ is Fredholm and $\mathrm{ind}(S_1(\mu)) = \mathrm{ind}(S_1(\mu) \oplus I_Z)$.

(b) Suppose that $S_1(\lambda) \sim_\mu S_2(\lambda)$. If $S_1(\lambda)$ is finitely meromorphic at μ then $S_2(\lambda)$ is finitely meromorphic at μ. Moreover, $K := \int_\Gamma S_1'(\lambda) S_1(\lambda)^{-1} d\lambda$ is a finite rank operator and $\mathrm{Tr}\left(\int_\Gamma S_1'(\lambda) S_1(\lambda)^{-1} d\lambda\right) = \mathrm{Tr}\left(\int_\Gamma S_2'(\lambda) S_2(\lambda)^{-1} d\lambda\right)$, where Γ lies in a neighbourhood of μ enclosing μ and $S_1'(\lambda)$ is the derivative of $S_1(\lambda)$.

(c) Suppose that $S_1(\lambda)$ is Fredholm for $\lambda \in U$. If $S_1(\lambda) \oplus I_Z \sim_U S_2(\lambda) \oplus I_Z$ then $S_1(\lambda) \sim_U S_2(\lambda)$.

Proof. The proof of (a) is immediate. The proof of (b) follows from [23, Lemma XI.9.3] and the proof of [23, Theorem XI.9.1]. The result in (c) is proved in [27, Theorem 1.1]. □

For regular holomorphic Fredholm operator-valued functions, we have the following result; see [23, p.202-205] and [27].

Theorem 3.3.5. *[23, 28] Let $U \subset \mathbb{C}$ be open and connected. Let $T : U \longrightarrow L(X)$ be a regular holomorphic Fredholm operator-valued function.*

(a) Then the spectrum $\sigma(T)$ is at most countable and has no accumulation point inside U. Further, for $\mu \in \sigma(T)$ and $\lambda \in \rho(T)$ sufficiently close to μ, we have

$$T(\lambda)^{-1} = \sum_{n=-m}^{\infty} (\lambda - \mu)^n A_n,$$

where A_0 is a Fredholm operator with $\operatorname{ind}(A_0) = 0$ and $A_{-1}, \dots, A_{-m}$ are operators of finite rank. In particular, $T(\lambda)^{-1}$ is finitely meromorphic at μ.

(b) Let $\mu \in \sigma(T)$. There exist positive integers $m_1 \le m_2 \le \cdots \le m_r$ such that

$$T(\lambda) \sim_\mu D(\lambda) := P_0 + (\lambda - \mu)^{m_1} P_1 + \cdots + (\lambda - \mu)^{m_r} P_r,$$

where $P_0, P_1, \dots, P_r$ are mutually disjoint projections, $P_1, \dots, P_r$ have rank one, and $P_0 + P_1 + \cdots + P_r = I$. The positive integers $m_1 \le \cdots \le m_r$ are unique and are called partial multiplicities of μ. Further, $\dim(N(T(\mu))) = \operatorname{rank}(I - P_0) = r$.

The next result characterizes discrete eigenvalues of $T(\lambda)$.

Theorem 3.3.6. *Let $U \subset \mathbb{C}$ be open and let $T : U \longrightarrow L(X)$ be holomorphic and regular. Let $\mu \in \sigma(T)$ be an isolated point. Then the following are equivalent.*

(a) μ is a discrete eigenvalue of $T(\lambda)$.

(b) $T(\mu)$ is Fredholm and $\operatorname{ind}(T(\mu)) = 0$.

(c) There is a $\delta > 0$ such that $|z - \mu| < \delta \Longrightarrow T(z)$ is Fredholm.

(d) There exist positive integers $m_1 \le m_2 \le \cdots \le m_r$ such that

$$T(\lambda) \sim_\mu D(\lambda) := P_0 + (\lambda - \mu)^{m_1} P_1 + \cdots + (\lambda - \mu)^{m_r} P_r,$$

where $P_0, P_1, \dots, P_r$ are mutually disjoint projections, $P_1, \dots, P_r$ have rank one, and $P_0 + P_1 + \cdots + P_r = I$.

(e) $T(\lambda)^{-1}$ is finitely meromorphic at μ.

Proof. The assertions (a) $\Rightarrow$ (b) $\Rightarrow$ (c) are immediate. The assertion (c) $\Rightarrow$ (d) follows from Theorem 3.3.5(b). Note that $D(\lambda)^{-1}$ is finitely meromorphic at μ and $D(\lambda)^{-1} \sim_\mu T(\lambda)^{-1}$. Hence the assertion (d) $\Rightarrow$ (e) follows from Lemma 3.3.4. Finally, the assertion (e) $\Rightarrow$ (a) follows from Theorem 3.4.6, which is proved in section 3.4. □

We mention that the equivalence of (b) and (d) is proved in [23, Theorem XI.8.1]. It is shown in [28, Theorem 3.1] that (d) holds if and only if $T(\lambda)^{-1} = \sum_{j=-\nu}^{\infty} A_j(\lambda-\mu)^j$, where $A_j, j = -1, \ldots, -\nu$, are finite rank operators, A_0 is Fredholm and $\text{ind}(A_0) = 0$. In other words, (d) holds $\iff$ $T(\lambda)^{-1}$ is finitely meromorphic at μ, A_0 is Fredholm and $\text{ind}(A_0) = 0$. We show in Theorem 3.4.6 that the additional assumption on A_0 is redundant. In fact, (d) holds $\iff$ $T(\lambda)^{-1}$ is finitely meromorphic at μ.

Definition 3.3.7. *Let $T : U \longrightarrow L(X)$ be holomorphic and regular. Let $\mu \in \sigma_d(T)$. Then $T(\lambda) \sim_\mu D(\lambda) := P_0 + (\lambda-\mu)^{m_1}P_1 + \cdots + (\lambda-\mu)^{m_r}P_r$, where $D(\lambda)$ is as given in Theorem 3.3.6. The positive integers $m_1 \leq \cdots \leq m_r$ are called the partial multiplicities of μ. The algebraic multiplicity of μ is defined as $m(\mu, T) := m_1 + \cdots + m_r$. If $\sigma \subset \sigma_d(T)$ is a finite set then $m(\sigma, T) := \sum_{\mu\in\sigma} m(\mu, T)$ is the total algebraic multiplicity of the eigenvalues in σ.*

Notice that if $m_1 \leq \cdots \leq m_r$ are the partial multiplicities of $\mu \in \sigma_d(T)$ then $g(\mu, T) := \dim N(T(\mu)) = \text{rank}(I - P_0) = r$ is the geometric multiplicity of μ and m_r is the ascent of μ. The algebraic multiplicity $m(\mu, T)$ is also given by the logarithmic residue of $T(\lambda)$; see [23, 28]. The number of zeros (counting multiplicities) of a holomorphic function $f : U \longrightarrow \mathbb{C}$ inside a simple closed curve Γ is given by the logarithmic residue

$$\text{n}(\Gamma, f) = \frac{1}{2\pi i}\int_\Gamma f'(z)f(z)^{-1}dz,$$

where $f'(z)$ is the derivative of $f(z)$ with respect to z. An operator version of the logarithmic residue theorem is proved in [28, 23, p.206] which gives the algebraic multiplicity $m(\lambda, T)$:

$$m(\mu, T) = \text{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda\right), \tag{3.2}$$

where $T'(\lambda)$ is derivative of $T(\lambda)$ with respect to λ and Γ is a simple closed rectifiable curve in $\rho(T)$ enclosing μ and isolating μ from the rest of $\sigma(T)$. The equality in (3.2) follows from the fact that since $T(\lambda) \sim_\mu D(\lambda)$, by Lemma 3.3.4 we have

$$\text{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda\right) = \text{Tr}\left(\frac{1}{2\pi i}\int_\Gamma D'(\lambda)D(\lambda)^{-1}d\lambda\right) = m_1 + \cdots + m_r.$$

For the special case when $T(\lambda) := \lambda I - A$, it follows that

$$\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda = \frac{1}{2\pi i}\int_\Gamma (\lambda I - A)^{-1}d\lambda =: P_\mu$$

is the spectral projection of A corresponding to the eigenvalue μ. Hence we have $m(\mu, T) = \mathrm{Tr}(P_\mu) = \mathrm{rank}(P_\mu) = m(\mu, A)$. We end this section with the statement of an operator version of the Rouche's theorem; see [28][23, p.206].

Theorem 3.3.8. *[23, 28] Let $U \subset \mathbb{C}$ be open and connected. Let $T, S : U \longrightarrow L(X)$ be holomorphic such that $T(\lambda)$ is regular and Fredholm for all $\lambda \in U$. Let $\Gamma \subset \rho(T)$ be a simple closed rectifiable curve such that $\mathrm{Int}(\Gamma) \subset U$. Set $W(\lambda) := T(\lambda) + S(\lambda)$ for $\lambda \in U$. If $\max_{\lambda\in\Gamma} \|T(\lambda)^{-1}S(\lambda)\| < 1$ then $\Gamma \subset \rho(W)$ and $W(\lambda)$ is Fredholm for all $\lambda \in \mathrm{Int}(\Gamma)$. Let $\sigma := \sigma(T) \cap \mathrm{Int}(\Gamma)$ and $\tau := \sigma(W) \cap \mathrm{Int}(\Gamma)$. Then $m(\sigma, T) = m(\tau, W)$.*

Note that Theorem 3.3.8 holds when $T : U \longrightarrow L(X)$ is holomorphic and regular, and $T(\lambda)$ is Fredholm for all $\lambda \in \mathrm{Int}(\Gamma)$; see [23, Theorem XI.9.1]. In particular, Theorem 3.3.8 holds when σ consists of discrete eigenvalues $T(\lambda)$.

3.4 Linearization of holomorphic operator-valued functions

Linearization is a technique that transforms a nonlinear eigenvalue problem to a linear eigenvalue problem. In this section, we consider linearization of holomorphic operator-valued functions which we utilize to derive perturbation bounds for eigenvalues. Linearization of a holomorphic operator-valued function is defined as follows.

Definition 3.4.1. *Let $\Omega \subset \mathbb{C}$ be a simply connected bounded domain. Let X and $\mathbb{X}$ be Banach spaces. Let $T : \Omega \longrightarrow L(X)$ be holomorphic and regular. Then an operator $\mathbb{T} \in L(\mathbb{X})$ is said to be a linearization of $T(\lambda)$ on Ω if there exists a Banach space Z such that*

$$\mathbb{T} - \lambda I \sim_\Omega T(\lambda) \oplus I_Z, \tag{3.3}$$

that is, there exist holomorphic and invertible operator functions $E : \Omega \to L(X \oplus Z, \mathbb{X})$ and $F : \Omega \to L(\mathbb{X}, X \oplus Z)$ such that $F(\lambda)(\lambda I - \mathbb{T})E(\lambda) = T(\lambda) \oplus I_Z$ for all $\lambda \in \Omega$, where I is the identity operator on $\mathbb{X}$ and I_Z is the identity operator on Z.

Observe that $\sigma(T) = \sigma(\mathbb{T}) \cap \Omega$. Also, if $T(\lambda)$ is Fredholm then so is $\mathbb{T} - \lambda I$. Hence if μ is a discrete eigenvalue of $T(\lambda)$ then μ is a discrete eigenvalue of $\mathbb{T}$.

Theorem 3.4.2. *Let $T : \Omega \longrightarrow L(X)$ be regular and holomorphic. Let $\mathbb{T} \in L(\mathbb{X})$ be a linearization of $T(\lambda)$ on Ω. Then $\sigma(T) = \sigma(\mathbb{T}) \cap \Omega$. Let $\mu \in \sigma(T)$ be an isolated point* *and* $\Gamma \subset \rho(T)$ *be such that* $\sigma(T) \cap \mathrm{Int}(\Gamma) = \{\mu\}$ *and* $\mathrm{Int}(\Gamma) \subset \Omega$. *Then* $\Gamma \subset \rho(\mathbb{T})$ *and* $\mu \in \sigma_d(T) \iff \mu \in \sigma_d(\mathbb{T})$. *Suppose that* $\mu \in \sigma_d(T)$.

(a) Let $\mathbb{P} := \frac{1}{2\pi i}\int_\Gamma (\lambda I - \mathbb{T})^{-1} d\lambda$ *be the spectral projection associated with* $\mathbb{T}$ *and* μ. *Then we have*

$$m(\mu, T) = \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda\right) = \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma (\lambda I - \mathbb{T})^{-1}d\lambda\right) = \mathrm{rank}(\mathbb{P}),$$

that is, the algebraic multiplicity of μ *as an eigenvalue of* $T(\lambda)$ *is the same as the algebraic multiplicity of* μ *as an eigenvalue of* $\mathbb{T}$.

(b) Moreover, if $T(\lambda) \sim_\mu D(\lambda)$ *then* $\mathbb{T} - \lambda I \sim_\mu D(\lambda) \oplus I_Z$, *where* $D(\lambda)$ *is as given in Theorem 3.3.6. Hence the partial multiplicities of* μ *as an eigenvalue of* $T(\lambda)$ *are the same as the partial multiplicities of* μ *as an eigenvalue of* $\mathbb{T}$. *Also, the ascent of* μ *as an eigenvalue of* $T(\lambda)$ *is the same as the ascent of* μ *as an eigenvalue of* $\mathbb{T}$.

Proof. Let $\lambda \in \Omega$. Then it follows from (3.3) that $T(\lambda)$ is invertible $\iff$ $\mathbb{T} - \lambda I$ is invertible. Hence $\sigma(T) = \sigma(\mathbb{T}) \cap \Omega$ and $\Gamma \subset \rho(\mathbb{T})$. By Lemma 3.3.4, $T(\mu)$ is Fredholm $\iff$ $\mathbb{T} - \mu I$ is Fredholm. Hence $\mu \in \sigma_d(T) \iff \mu \in \sigma_d(\mathbb{T})$.

Set $W(\lambda) := T(\lambda) \oplus I_Z$. Then $W'(\lambda)W(\lambda)^{-1} = T'(\lambda)T(\lambda)^{-1} \oplus 0_Z$, where 0_Z is the zero operator on Z. Since $W(\lambda) \sim_\mu \mathbb{T} - \lambda I$, by Lemma 3.3.4 we have

$$\begin{aligned} m(\mu, T) &= \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda\right) = \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma W'(\lambda)W(\lambda)^{-1}d\lambda\right) \\ &= \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma (\lambda I - \mathbb{T})^{-1}d\lambda\right) = \mathrm{rank}(\mathbb{P}). \end{aligned}$$

This proves (a).

Note that $T(\lambda) \oplus I_Z = F(\lambda)(\mathbb{T} - \lambda I)E(\lambda)$ with $E(\lambda)$ and $F(\lambda)$ being holomorphic and invertible for all $\lambda \in \Omega$. Hence if $T(\lambda) \sim_\mu D(\lambda)$ then $\mathbb{T} - \lambda I \sim_\mu T(\lambda) \oplus I_Z \sim_\mu D(\lambda) \oplus I_Z$. This shows that the partial multiplicities of μ as the eigenvalue of $T(\lambda)$ are the same as the partial multiplicities of μ as the eigenvalue of $\mathbb{T}$. It also follows that the orders of the pole of $T(\lambda)^{-1}$ and $(\mathbb{T} - \lambda I)^{-1}$ at μ are the same. In fact, m_r is the ascent of μ. This proves (b). □

Next, we describe construction of a linearization of $T(\lambda)$ due to Gohberg-Kaashoek-Lay; see [27, 23, p.38]. For the rest of the chapter, we assume that Ω is a simply connected bounded open subset of $\mathbb{C}$ such that the boundary $\partial\Omega$ is a simple closed rectifiable curve and is oriented positively. Let $C(\partial\Omega, X)$ denote the Banach space of all X-valued continuous functions on Ω endowed with the supremum norm, that is,

$$C(\partial\Omega, X) := \{f : \partial\Omega \longrightarrow X \mid f \text{ is continuous}\} \text{ and } \|f\|_\infty := \sup_{z\in\partial\Omega} \|f(z)\|.$$

We denote by $\mathbb{H}(\Omega, L(X))$ the vector space of $L(X)$-valued holomorphic functions on Ω which are continuous on the closure $\bar{\Omega}$, that is,

$$\mathbb{H}(\Omega, L(X)) := \{T : \Omega \longrightarrow L(X) \mid T \text{ is holomorphic on } \Omega \text{ and continuous on } \bar{\Omega}\}.$$

Let $T \in \mathbb{H}(\Omega, L(X))$ be regular. The Gohberg-Kaashoek-Lay construction of a linearization of $T(\lambda)$ is as follows.

Theorem 3.4.3. *[27, 23] Let $T \in \mathbb{H}(\Omega, L(X))$. Define $\mathbb{T} : C(\partial\Omega, X) \longrightarrow C(\partial\Omega, X)$ by*

$$(\mathbb{T}f)(\lambda) := \lambda f(\lambda) - \frac{1}{2\pi i}\int_{\partial\Omega}(I - T(w))f(w)dw, \tag{3.4}$$

where I is the identity operator on X. Then $\mathbb{T}$ is a bounded operator and there exists a Banach space Z such that $T(\lambda) \oplus I_Z \sim_\Omega \mathbb{T} - \lambda I$, where I is the identity operator on $C(\partial\Omega, X)$. Hence $\mathbb{T}$ is a linearization of $T(\lambda)$ on Ω.

The proof of Theorem 3.4.3 in [27, 23] assumes that $0 \in \Omega$ and defines the Banach space Z by $Z := \{f \in C(\partial\Omega, X) \mid \int_{\partial\Omega} \frac{1}{z} f(z)dz = 0\}$ and then shows that $\mathbb{T} - \lambda I \sim_\Omega T(\lambda) \oplus I_Z$.

We drop the assumption that $0 \in \Omega$ as in general Ω cannot always be expected to contain 0. Let $\lambda_0 \in \Omega$ be arbitrary but fixed. Since Ω is open, $\lambda_0 \notin \partial\Omega$. Now define

$$Z := \{f \in C(\partial\Omega, X) \mid \int_{\partial\Omega} \frac{1}{z - \lambda_0} f(z)dz = 0\}. \tag{3.5}$$

Then Z is a Banach space and an appropriate modification of the proof of Theorem 3.4.3 in [27, 23] shows that $\mathbb{T} - \lambda I \sim_\Omega T(\lambda) \oplus I_Z$. For completeness, we provide a complete proof.

Proof. Define the linear map $\mathcal{W} : C(\partial\Omega, X) \to X$ by

$$\mathcal{W}f := \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{z - \lambda_0} f(z)dz.$$

We define the canonical embedding $\mathcal{I} : X \to C(\partial\Omega, X), x \longmapsto \mathcal{I}x$, by

$$(\mathcal{I}x)(w) := x \quad \text{for} \quad w \in \partial\Omega.$$

and identify X as a subspace of $C(\partial\Omega, X)$, that is, x is a constant function in $C(\partial\Omega, X)$. Then it follows that $\mathcal{W}\mathcal{I} = I$ and $P := \mathcal{I}\mathcal{W} : C(\partial\Omega, X) \to C(\partial\Omega, X)$ is a projection. Now note that the Banach space Z defined in (3.5) is given by $Z = N(P)$.

Next, we define the canonical isomorphism $J : X \oplus Z \to C(\partial\Omega, X)$ given by

$$J(x, y) := \mathcal{I}x + y \quad \text{for } (x, y) \in X \oplus Z.$$

and identify $X \oplus Z$ with $C(\partial\Omega, X)$. Note that $J^{-1}f = (\mathcal{W}f, (I-P)f)$ for $f \in C(\partial\Omega, X)$. Now define $V : C(\partial\Omega, X) \to C(\partial\Omega, X), f \longmapsto Vf$, by

$$(Vf)(w) := wf(w) \quad \text{for } w \in \partial\Omega.$$

Note that, for $\lambda \in \Omega$, $\lambda I - V$ is invertible and $(\lambda I - V)^{-1}f(w) = (\lambda - w)^{-1}f(w)$ for $w \in \partial\Omega$ and $f \in C(\partial\Omega, X)$. Next, define $M : C(\partial\Omega, X) \to C(\partial\Omega, X), f \longmapsto Mf$, by

$$(Mf)(w) := T(w)f(w) \quad \text{for } w \in \partial\Omega.$$

Then $MV = VM$ and hence M and $(\lambda I - V)^{-1}$ commute. For $\lambda \in \Omega$, define

$$B(\lambda) := I + P(V - \lambda_0 I)(\lambda I - V)^{-1} - P(V - \lambda_0)(\lambda I - V)^{-1}M,$$

$E(\lambda) := (\lambda I - V)^{-1}J$ and $F(\lambda) := J^{-1}(I - PB(\lambda)(I-P))$. Since $P = \mathcal{I}\mathcal{W}$ is a projection, it follows that $(I - PB(\lambda)(I-P))^{-1} = I + PB(\lambda)(I-P)$ for $\lambda \in \Omega$. This shows that $E(\lambda)$ and $F(\lambda)$ are invertible for $\lambda \in \Omega$.

Now by (3.4), for $f \in C(\partial\Omega, X)$, we have

$$\begin{aligned}
(\mathbb{T}f)(z) &= zf(z) - \frac{1}{2\pi i}\int_{\partial\Omega}(I - T(w))f(w)dw \\
&= zf(z) - \frac{1}{2\pi i}\int_{\partial\Omega}\frac{(w-\lambda_0)}{w-\lambda_0}(I - T(w))f(w)dw \\
&= zf(z) - \frac{1}{2\pi i}\int_{\partial\Omega}\frac{wf(w)}{w-\lambda_0}dw + \frac{1}{2\pi i}\int_{\partial\Omega}\frac{wT(w)f(w)}{w-\lambda_0}dw \\
&\quad + \frac{1}{2\pi i}\int_{\partial\Omega}\frac{\lambda_0 f(w)}{w-\lambda_0}dw - \frac{1}{2\pi i}\int_{\partial\Omega}\frac{\lambda_0 T(w)f(w)}{w-\lambda_0}dw \\
&= zf(z) - \frac{1}{2\pi i}\int_{\partial\Omega}\frac{(Vf)(w)}{w-\lambda_0}dw + \frac{1}{2\pi i}\int_{\partial\Omega}\frac{(VMf)(w)}{w-\lambda_0}dw \\
&\quad + \frac{1}{2\pi i}\int_{\partial\Omega}\frac{\lambda_0 f(w)}{w-\lambda_0}dw - \frac{1}{2\pi i}\int_{\partial\Omega}\frac{\lambda_0 Mf(w)}{w-\lambda_0}dw \\
&= (Vf)(z) - (\mathcal{I}\mathcal{W}Vf)(z) + (\mathcal{I}\mathcal{W}VMf)(z) + \lambda_0(\mathcal{I}\mathcal{W}f)(z) - \lambda_0(\mathcal{I}\mathcal{W}Mf)(z) \\
&= ((V - PV + \lambda_0 P + PVM - \lambda_0 PM)f)(z)
\end{aligned}$$

for all $z \in \partial\Omega$. Thus we have $\mathbb{T} = V - PV + \lambda_0 P + PVM - \lambda_0 PM$.

For all $\lambda \in \Omega$, we have

$$\begin{aligned}
F(\lambda)(\lambda I - \mathbb{T})E(\lambda) &= F(\lambda)(\lambda I - V + PV - \lambda_0 P - PVM + \lambda_0 PM)(\lambda I - V)^{-1}J \\
&= F(\lambda)(I + P(V - \lambda_0)(\lambda I - V)^{-1} - P(V - \lambda_0)(\lambda I - V)^{-1}M)J \\
&= J^{-1}(I - PB(\lambda)(I-P))B(\lambda)J = J^{-1}PB(\lambda)PJ + J^{-1}(I-P)J.
\end{aligned}$$

Now for $\lambda \in \Omega$ and $(x, y) \in X \oplus Z$, we have

$$J^{-1}PB(\lambda)PJ(x,y) = J^{-1}PB(\lambda)P(\mathcal{I}x + y) = J^{-1}PB(\lambda)P\mathcal{I}x$$
$$= (\mathcal{W}PB(\lambda)\mathcal{I}x, (I-P)PB(\lambda)\mathcal{I}x) = (\mathcal{W}B(\lambda)\mathcal{I}x, 0).$$

Next, for $\lambda \in \Omega$ and $x \in X$, we have

$$\begin{aligned}
\mathcal{W}B(\lambda)\mathcal{I}x &= \mathcal{W}[I + P(V-\lambda_0)(\lambda I - V)^{-1} - P(V-\lambda_0 I)(\lambda I - V)^{-1}M]\mathcal{I}x \\
&= \mathcal{W}\mathcal{I}x + \mathcal{W}(V-\lambda_0)(\lambda I - V)^{-1}\mathcal{I}x - \mathcal{W}(V-\lambda_0 I)(\lambda I - V)^{-1}M\mathcal{I}x \\
&= x + \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{w-\lambda_0}((V-\lambda_0 I)(\lambda I - V)^{-1}\mathcal{I}x)(w)dw \\
&- \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{w-\lambda_0}((V-\lambda_0 I)(\lambda I - V)^{-1}M\mathcal{I}x)(w)dw \\
&= x + \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{\lambda - w}xdw - \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{\lambda - w}T(w)xdw \\
&= x + (-1)x + T(\lambda)x = T(\lambda)x.
\end{aligned}$$

This shows that $J^{-1}PB(z)PJ = \begin{bmatrix} T(\lambda) & 0 \\ 0 & 0 \end{bmatrix}$ for $\lambda \in \Omega$. Since $J^{-1}(I-P)J = \begin{bmatrix} 0 & 0 \\ 0 & I_Z \end{bmatrix}$, we have

$$\begin{aligned}
F(\lambda)(\lambda I - \mathbb{T})E(\lambda) &= J^{-1}PB(\lambda)PJ + J^{-1}(I-P)J \\
&= \begin{bmatrix} T(\lambda) & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & I_Z \end{bmatrix} = \begin{bmatrix} T(\lambda) & 0 \\ 0 & I_Z \end{bmatrix}
\end{aligned}$$

for all $\lambda \in \Omega$. This completes the proof. □

Theorem 3.4.4. *[27, Theorem 2.3] Let $T \in \mathbb{H}(\Omega, L(X))$ be regular and $\mathbb{T} : C(\partial\Omega)^n \longrightarrow C(\partial\Omega)^n$ be a linearization of $T(\lambda)$ as given in Theorem 3.4.3. Then $\sigma(\mathbb{T}) = \sigma(T) \cup \partial\Omega$.*

The next result describes the bijective correspondence between eigenvectors of $T(\lambda)$ and $\mathbb{T}$. This correspondence allows us to recover eigenvectors of $T(\lambda)$ from those of $\mathbb{T}$.

Proposition 3.4.5. *Let $T \in \mathbb{H}(\Omega, L(X))$ be regular and $\mu \in \sigma(T)$ be an eigenvalue. Let $\mathbb{T}$ be a linearization of $T(\lambda)$ as given in Theorem 3.4.3. Then the linear maps $\mathcal{E} : N(T(\mu)) \longrightarrow N(\mathbb{T} - \mu I), x \longmapsto \mathcal{E}x$, and $\mathcal{F} : N(\mathbb{T} - \mu I) \longrightarrow N(T(\mu))$ given by*

$$(\mathcal{E}x)(w) := \frac{x}{\mu - w}, \; w \in \partial\Omega, \; \text{and } \mathcal{F}f := \frac{1}{2\pi i}\int_{\partial\Omega} \frac{(\mu - w)}{w - \lambda_0} f(w)dw$$

are isomorphisms, where $\lambda_0 \in \Omega$ is fixed as in (3.5).

Proof. We have $F(\lambda)(\lambda I - \mathbb{T})E(\lambda) = T(\lambda) \oplus I_Z$ for all $\lambda \in \Omega$, where $E(\lambda)$ and $F(\lambda)$ are as defined in the proof of Theorem 3.4.3. Let $x \in X$. Then it follows that $x \in N(T(\mu)) \iff E(\mu)(x,0) \in N(\mathbb{T} - \mu I)$. Now, for $x \in N(T(\mu))$, we have

$$\mathcal{E}x = (\mu I - V)^{-1}\mathcal{I}x = (\mu I - V)^{-1}J(x,0) = E(\mu)(x,0)$$

which shows that $\mathcal{E}x \in N(\mathbb{T} - \mu I)$. Hence $\mathcal{E}$ is an isomorphism.

Let $f \in C(\partial\Omega, X)$. It follows that $f \in N(\mathbb{T} - \mu I) \iff E(\mu)^{-1}f \in N(T(\mu) \oplus I_Z)$. Now $E(\mu)^{-1}f = (J^{-1}(\mu I - V))f = (\mathcal{W}(\mu I - V))f, (I - P)(\mu I - V)f)$ shows that $E(\mu)^{-1}f \in N(T(\mu) \oplus I_Z) \iff \mathcal{W}(\mu I - V))f \in N(T(\mu))$ and $(I - P)(\mu I - V)f = 0$.

Note that

$$\begin{aligned}\mathcal{W}(\mu I - V))f &= \frac{1}{2\pi i}\int_\Gamma \frac{1}{w - \lambda_0}(\mu I - V)f(w)dw \\ &= \frac{1}{2\pi i}\int_\Gamma \frac{1}{w - \lambda_0}(\mu I - w)f(w)dw = \mathcal{F}f.\end{aligned}$$

This shows that $f \in N(\mathbb{T} - \mu I) \Leftrightarrow \mathcal{F}f \in N(T(\mu))$ and that $\mathcal{F}$ is an isomorphism. □

We now prove that μ is a discrete eigenvalue of $T(\lambda) \iff T(\lambda)^{-1}$ is finitely meromorphic at μ.

Theorem 3.4.6. *Let $T \in \mathbb{H}(\Omega, L(X))$ be regular and $\mu \in \sigma(T)$. Then $\mu \in \sigma_d(T) \iff T(\lambda)^{-1}$ is finitely meromorphic at μ. Moreover, if $T(\lambda)^{-1}$ is finitely metamorphic at μ, that is, $T(\lambda)^{-1} = \sum_{j=-\nu}^{\infty} A_j(\lambda-\mu)^j$, where $A_j, j = -1, \ldots, -\nu$, are finite rank operators, then A_0 is Fredholm and* $\mathrm{ind}(A_0) = 0$.

Proof. If $\mu \in \sigma_d(T)$ then as shown in the proof of the assertions $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e)$ in Theorem 3.3.6, $T(\lambda)^{-1}$ is finitely meromorphic at μ.

Conversely, suppose that $T(\lambda)^{-1}$ is finitely meromorphic at μ. Then μ is an isolated point of $\sigma(T)$. Since $T(\lambda)^{-1}$ is finitely meromorphic at μ, $T(\lambda)^{-1} \oplus I_Z$ is finitely meromorphic at μ. Now, consider the linearization $\mathbb{T}$ of $T(\lambda)$ on Ω as given in (3.4). Since $T(\lambda) \oplus I_Z \sim_\Omega \mathbb{T} - \lambda I$, by Lemma 3.3.4, $\mathbb{T} - \lambda I$ is finitely meromorphic at μ. Hence by Theorem 3.2.3, $\mu \in \sigma_d(\mathbb{T})$. Now by Theorem 3.4.2, $\mu \in \sigma_d(T)$.

By (4.1), the constant term in the Laurent series expansion of $(\mathbb{T}-\lambda I)^{-1}$ at μ is given by the reduced resolvent $\mathbb{S}$ of $\mathbb{T} - \mu I$. Since $T(\lambda) \oplus I_Z = F(\lambda)(\mathbb{T} - \lambda I)E(\lambda)$, it follows that $A_0 \oplus I_Z = E(\lambda)^{-1}\mathbb{S}F(\lambda)^{-1}$ which shows that $A_0 \oplus I_Z \sim_\Omega \mathbb{S}$. By Remark 5.2.9, $\mathbb{S}$ is Fredholm and $\mathrm{ind}(\mathbb{S}) = 0$. Hence by Lemma 3.3.4, A_0 is Fredholm and $\mathrm{ind}(A_0) = \mathrm{ind}(\mathbb{S}) = 0$. This completes the proof. □

It is shown in [28, Theorem 3.1] that $T(\lambda) \sim_\mu D(\lambda) \iff T(\lambda)^{-1}$ is finitely meromorphic, A_0 is Fredholm and $\mathrm{ind}(A_0) = 0$, where $D(\lambda)$ is as in Theorem 3.3.6 and A_0 as in Theorem 3.4.6. Theorem 3.4.6 shows that the assumption on A_0 is redundant. Thus we have the following.

Corollary 3.4.7. *Let $T \in \mathbb{H}(\Omega, L(X))$ be regular and $\mu \in \Omega$. Then $T(\lambda)^{-1}$ is finitely meromorphic $\iff$ $T(\lambda)^{-1} = \sum_{j=-\nu}^{\infty} A_j(\lambda-\mu)^j$, where $A_j, j = -1, \ldots, -\nu$, are finite rank operators, A_0 is Fredholm and* $\operatorname{ind}(A_0) = 0$.

3.5 Perturbation theory for discrete eigenvalues

Let $T, V \in \mathbb{H}(\Omega, L(X))$ and $t \in \mathbb{C}$. Suppose that $T(\lambda)$ is regular. We now study the effect of the perturbation $T(\lambda)+tV(\lambda)$ on the eigenvalues of $T(\lambda)$ when t varies in $\mathbb{C}$. Let $r_\sigma(A)$ denote the spectral radius of an operator $A \in L(X)$. Since $T(\lambda)$ is holomorphic, it is well known [7] that the map $\Omega \to \mathbb{R}$, $\lambda \mapsto r_\sigma(T(\lambda))$, is a subharmonic function on Ω. Hence $r_\sigma(T(\lambda))$ attains its maximum on the boundary $\partial\Omega$.

Let $\Gamma \subset \rho(T)$ be a simple closed positively oriented rectifiable curve. Define

$$\partial_\Gamma := \{t \in \mathbb{C} : |t| \max_{z\in\Gamma} r_\sigma(V(z)T(z)^{-1}) < 1\}.$$

Recall that $\operatorname{Int}(\Gamma)$ is the interior of the region enclosed by the curve Γ.

Theorem 3.5.1. *Define $W(\lambda,t) := T(\lambda) + tV(\lambda)$ for $\lambda \in \Omega$ and $t \in \mathbb{C}$. Suppose that $\Gamma \subset \rho(T)$. Then the following results hold:*

(a) We have $\Gamma \subset \rho(T + tV)$ for $t \in \partial_\Gamma$.

(b) The map $\partial_\Gamma \longrightarrow L(X)$, $t \longmapsto W(\lambda,t)^{-1}$, is holomorphic for each fixed $\lambda \in \Gamma$.

(c) The map $\Gamma \longrightarrow L(X)$, $\lambda \longmapsto W(\lambda,t)^{-1}$, is holomorphic for each fixed $t \in \partial_\Gamma$.

Proof. Let $z \in \Gamma$. Then for all $t \in \partial_\Gamma$, $r_\sigma(tV(z)T(z)^{-1}) < 1 \Rightarrow (I + tV(z)T(z)^{-1})^{-1}$ exists. This shows that $T(z) + tV(z) = (I + tV(z)T(z)^{-1})T(z)$ is invertible for all $t \in \partial_\Gamma$. Hence $\Gamma \subset \rho(T + tV)$ for all $t \in \partial_\Gamma$.

Let $GL(X)$ denote the group of invertible operators in $L(X)$. Then $GL(X)$ is open and the map $GL(X) \longrightarrow GL(X), A \longmapsto A^{-1}$, is smooth. Fix $\lambda \in \Gamma$. Since $W(\lambda,t) \in GL(X)$ for $t \in \partial_\Gamma$ and $t \longmapsto W(\lambda,t)$ is holomorphic on ∂_Γ, by the chain rule the function $t \longmapsto W(\lambda,t)^{-1}$ is holomorphic on ∂_Γ.

Alternatively, set $\delta := \left(\max_{z\in\Gamma} r_\sigma(V(z)W(z,t_0,)^{-1})\right)^{-1}$. Let $t \in \partial_\Gamma$ be such that $|t-t_0| < \delta$. Fix $\lambda \in \Gamma$. Then we have $|t-t_0| r_\sigma(V(\lambda)W(\lambda,t_0)^{-1}) < 1 \Longrightarrow I + (t-t_0)V(\lambda)W(\lambda,t_0)^{-1}$ is invertible and

$$\begin{aligned} W(\lambda,t)^{-1} &= W(\lambda,t_0)^{-1}(I + (t-t_0)V(\lambda)W(\lambda,t_0)^{-1})^{-1} \\ &= W(\lambda,t_0)^{-1}\sum_{k=0}^{\infty}(-1)^k(V(\lambda)W(\lambda,t_0)^{-1})^k(t-t_0)^k. \end{aligned}$$

Finally, fix $t \in \partial_\Gamma$. By part (a), $W(t,\lambda) \in GL(X)$ for $\lambda \in \Gamma$. Since $\lambda \longmapsto W(\lambda, t)$ is holomorphic on Γ and $A \longmapsto A^{-1}$ is smooth on $GL(X)$, by the chain rule the function $\lambda \longmapsto W(\lambda, t)^{-1}$ is holomorphic on Γ. $\square$

Now we consider the linearization of $W(\lambda, t)$ on Ω as defined in (3.4). For $t \in \mathbb{C}$, define the linearization $\mathbb{T}(t) : C(\partial\Omega, X) \longrightarrow C(\partial\Omega, X)$ of $W(\lambda, t)$ by

$$(\mathbb{T}(t)f)(\lambda) := \lambda f(\lambda) - \frac{1}{2\pi i}\int_{\partial\Omega}(I - W(w,t))f(w)dw, \tag{3.6}$$

where I is the identity operator on X. Note that $\mathbb{T} = \mathbb{T}(0)$ is the linearization of $T(\lambda)$.

Theorem 3.5.2. *We have $\Gamma \subset \rho(\mathbb{T}(t))$ for all $t \in \partial_\Gamma$. Define the spectral projection*

$$\mathbb{P}(t) := \frac{1}{2\pi i}\int_\Gamma (zI - \mathbb{T}(t))^{-1}dz$$

associated with $\mathbb{T}(t)$ and Γ. Then $t \longmapsto \mathbb{P}(t)$ is holomorphic on ∂_Γ. Further, we have $\mathrm{rank}(\mathbb{P}(t)) = \mathrm{rank}(\mathbb{P})$ *for all $t \in \partial_\Gamma$, where $\mathbb{P} = \mathbb{P}(0)$ is the spectral projection associated with $\mathbb{T} = \mathbb{T}(0)$ and Γ.*

(a) Let $\sigma_\Gamma := \sigma(T) \cap \mathrm{Int}(\Gamma)$ be nonempty. Then $\sigma_\Gamma = \sigma(\mathbb{T}) \cap \mathrm{Int}(\Gamma)$ and

$$\sigma_\Gamma(t) := \sigma(T + tV) \cap \mathrm{Int}(\Gamma) = \sigma(\mathbb{T}(t)) \cap \mathrm{Int}(\Gamma) \neq \emptyset \quad \textit{for all} \ \ t \in \partial_\Gamma.$$

(b) Suppose that $\mathrm{rank}(\mathbb{P}) = \ell$. *Then σ_Γ consists of at most ℓ discrete eigenvalues of $T(\lambda)$ of total algebraic multiplicity ℓ. Further, $\sigma_\Gamma(t)$ consists of at most ℓ discrete eigenvalues of $W(\lambda, t)$ of total algebraic multiplicity ℓ for all $t \in \partial_\Gamma$. Further, we have*

$$\begin{aligned} m(\sigma_\Gamma(t), T + tV) &= \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma \frac{\partial W(\lambda,t)}{\partial \lambda} W(\lambda,t)^{-1}d\lambda\right) \\ &= \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda)T(\lambda)^{-1}d\lambda\right) = m(\sigma_\Gamma, T) \end{aligned}$$

for all $t \in \partial_\Gamma$.

(c) Let $\mu_1(t), \ldots, \mu_\ell(t)$ be the ℓ discrete eigenvalues in $\sigma_\Gamma(t)$ counted according to their algebraic multiplicities. Let $\mu_{\mathrm{av}}(t) := \frac{\mu_1(t)+\cdots+\mu_\ell(t)}{\ell}$ be the arithmetic mean of the eigenvalues. Then $t \longmapsto \mu_{\mathrm{av}}(t)$ is holomorphic on ∂_Γ. In particular, if $\sigma_\Gamma(t)$ consists of a simple eigenvalue $\mu(t)$ then $t \longmapsto \mu(t)$ is holomorphic on ∂_Γ.

Proof. By Theorem 3.4.3, we have $W(\lambda, t) \oplus I_Z \sim_\Omega \mathbb{T}(t) - \lambda I$ and by Theorem 3.5.1, we have $\Gamma \subset \rho(T + tV)$ for $t \in \partial_\Gamma$. Hence $\Gamma \subset \rho(\mathbb{T}(t))$ for all $t \in \partial_\Gamma$. Thus for each $\lambda \in \Gamma$, $\mathbb{T}(t) - \lambda I \in GL(C(\partial\Omega, X))$ for all $t \in \partial_\Gamma$. It follows from (3.6) that $t \longmapsto \mathbb{T}(t)$ is holomorphic on ∂_Γ for each fixed $\lambda \in \Gamma$. Since $A \longmapsto A^{-1}$ is a smooth map on

$GL(C(\partial\Omega, X))$, by the chain rule we conclude that $t \longmapsto (\lambda I - \mathbb{T}(t))^{-1}$ is holomorphic on ∂_Γ for each fixed $\lambda \in \Gamma$.

Let $t_0 \in \partial_\Gamma$. Then there exists a $\delta > 0$ such that for $t \in \partial_\Gamma$ and $|t - t_0| < \delta$, we have the power series $(\lambda I - \mathbb{T}(t))^{-1} = \sum_{n=0}^{\infty} A_n(\lambda)(t - t_0)^n$ for each fixed $\lambda \in \Gamma$ and the series converges absolutely and uniformly on Γ. Hence we have

$$\mathbb{P}(t) := \frac{1}{2\pi i}\int_\Gamma (\lambda I - \mathbb{T}(t))^{-1} d\lambda = \sum_{n=0}^{\infty} \frac{1}{2\pi i}\int_\Gamma (t - t_0)^n A_n(\lambda) d\lambda = \sum_{n=0}^{\infty} (t - t_0)^n P_n$$

which shows that $\mathbb{T}(t)$ is holomorphic at t_0. Since $t_0 \in \partial_\Gamma$ is arbitrary, $t \longmapsto \mathbb{P}(t)$ is holomorphic on ∂_Γ. As ∂_Γ is connected and $\mathbb{P}(t)$ is holomorphic on ∂_Γ, we conclude that $\mathrm{rank}(\mathbb{P}(t))$ is constant for all $t \in \partial_\Gamma$. Hence $\mathrm{rank}(\mathbb{P}(t)) = \mathrm{rank}(\mathbb{P}(0)) = \mathrm{rank}(\mathbb{P})$ for all $t \in \partial_\Gamma$.

(a) Since $\mathbb{T}$ is a linearization of $T(\lambda)$, we have $\sigma(T) = \sigma(\mathbb{T}) \cap \Omega$. Hence we have $\sigma_\Gamma = \sigma(\mathbb{T}) \cap \mathrm{Int}(\Gamma) \neq \emptyset$. Consequently $\mathbb{P} \neq 0$. Again, since $\mathbb{T}(t)$ is a linearization of $W(\lambda, t)$, we have $\sigma(T + tV) = \sigma(\mathbb{T}(t)) \cap \Omega$ for $t \in \mathbb{C}$. Hence $\sigma_\Gamma(t) = \sigma(\mathbb{T}(t)) \cap \mathrm{Int}(\Gamma)$ for $t \in \partial_\Gamma$.

Now, if possible, suppose that $\sigma_\Gamma(t_0) = \emptyset$ for some $t_0 \in \partial_\Gamma$. Then $\mathrm{Int}(\Gamma) \subset \rho(\mathbb{T}(t_0))$ and hence $\lambda \longmapsto (\lambda I - \mathbb{T}(t_0))^{-1}$ is holomorphic on $\mathrm{Int}(\Gamma)$. Hence by Cauchy's theorem $\mathbb{P}(t_0) = 0$ which contradicts that $\mathrm{rank}(\mathbb{P}(t_0)) = \mathrm{rank}(\mathbb{P}) \neq 0$. Hence $\sigma_\Gamma(t) \neq \emptyset$ for all $t \in \partial_\Gamma$.

(b) Since $\mathrm{rank}(\mathbb{P}) = \ell$, σ_Γ consists of at most ℓ discrete eigenvalues of $\mathbb{T}$ of total algebraic multiplicity ℓ. As $\mathbb{T}$ is a linearization of $T(\lambda)$, by Theorem 3.4.2 each eigenvalue of $\mathbb{T}$ in σ_Γ is a discrete eigenvalue of $T(\lambda)$ and $m(\mu, T) = m(\mu, \mathbb{T})$ for $\mu \in \sigma_\Gamma$. Since $\mathrm{rank}(\mathbb{P}) = \mathrm{rank}(\mathbb{P}(t)) = \ell$ for all $t \in \partial_\Gamma$, by Theorem 3.4.2, we have

$$\begin{aligned} m(\sigma_\Gamma(t), T + tV) &= \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma \frac{\partial W(\lambda, t)}{\partial \lambda} W(\lambda, t)^{-1} d\lambda\right) = \mathrm{rank}(\mathbb{P}(t)) \\ &= \mathrm{rank}(\mathbb{P}) = \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma T'(\lambda) T(\lambda)^{-1} d\lambda\right) = m(\sigma_\Gamma, T) \end{aligned}$$

for all $t \in \partial_\Gamma$.

(c) Note that $\mu_{\mathrm{av}}(t) = \frac{1}{\ell}\mathrm{Tr}(\mathbb{T}(t)\mathbb{P}(t))$. Since $t \longmapsto \mathbb{T}(t)$ and $t \longmapsto \mathbb{P}(t)$ are holomorphic on ∂_Γ, it follows that $t \longmapsto \mu_{\mathrm{av}}(t)$ is holomorphic on ∂_Γ. □

In view of Theorem 3.5.2, we have the following operator version of Rouche's theorem under a weaker assumption than that of Theorem 3.3.8.

Theorem 3.5.3 (Rouche's theorem)**.** *Let $T, V : \Omega \longrightarrow L(X)$ be holomorphic such that $T(\lambda)$ is regular. Set $W(\lambda) := T(\lambda) + V(\lambda)$ for $\lambda \in \Omega$. Let $\Gamma \subset \rho(T)$. If*

$$\max_{\lambda \in \Gamma} r_\sigma\left(V(\lambda)T(\lambda)^{-1}\right) < 1$$

then $\Gamma \subset \rho(W)$. *Let* $\sigma := \sigma(T) \cap \mathrm{Int}(\Gamma)$ *and* $\tau := \sigma(W) \cap \mathrm{Int}(\Gamma)$. *Suppose that* $\sigma \subset \sigma_d(T)$. *Then* $T(\lambda)$ *and* $W(\lambda)$ *are Fredholm operators of index zero for* $\lambda \in \Gamma \cup \mathrm{Int}(\Gamma)$ *and* $m(\sigma, T) = m(\tau, W)$.

Proof. By Theorem 3.5.2, $\Gamma \subset \rho(W)$, $\tau \subset \sigma_d(W)$ and $m(\sigma, T) = m(\tau, W)$. Note that $T(\lambda)$ is Fredholm for $\lambda \in \sigma$ and $T(\lambda)$ is invertible for $\lambda \in \mathrm{Int}(\Gamma) \setminus \sigma$. Hence $T(\lambda)$ is Fredholm with index zero for all $\lambda \in \Gamma \cup \mathrm{Int}(\Gamma)$. Similarly, $W(\lambda)$ is Fredholm with index zero for all $\lambda \in \Gamma \cup \mathrm{Int}(\Gamma)$. □

We now illustrate that the assumption in Theorem 3.5.3 is weaker than that in Theorem 3.3.8 by considering an example.

Example 3.5.4. *Let* $0 < \epsilon < 1$. *Let* $M_2(\mathbb{C})$ *be the Banach space of all* 2×2 *complex matrices equipped with the Hölder norm* $\|\cdot\|_\infty$. *Define* $T, V : \mathbb{C} \longrightarrow M_2(\mathbb{C})$ *by*

$$T(\lambda) := \begin{bmatrix} 4 & 41\lambda^3 + 3 \\ 0 & 2\lambda^2 \end{bmatrix} \text{ and } V(\lambda) := \begin{bmatrix} -\lambda^2 & 0 \\ 0 & -(\lambda^2 + \epsilon^2) \end{bmatrix}.$$

Then $\mu := 0$ *is the only eigenvalue of* T *of algebraic multiplicity* 2. *Hence* $\sigma(T) = \{0\}$. *On the other hand,* $\sigma(T + V) = \{\pm\epsilon, \pm 2\}$ *and each eigenvalue of* $T + V$ *is simple.*

Let Γ *be the unit circle* $|z| = 1$ *in* $\mathbb{C}$ *oriented positively. Then obviously* $\Gamma \subset \rho(T)$ *and* $\sigma(T) \cap \mathrm{Int}(\Gamma) = \{0\}$. *Now for* $\lambda \neq 0$, *we have*

$$T(\lambda)^{-1} = \begin{bmatrix} \frac{1}{4} & -\frac{(41\lambda^3+3)}{8\lambda^2} \\ 0 & \frac{1}{2\lambda^2} \end{bmatrix} \text{ and } V(\lambda)T(\lambda)^{-1} = \begin{bmatrix} -\frac{\lambda^2}{4} & \frac{(41\lambda^3+3)}{8\lambda^2} \\ 0 & -\frac{(\lambda^2+\epsilon^2)}{2\lambda^2} \end{bmatrix}.$$

Thus, we have $r_\sigma(V(\lambda)T(\lambda)^{-1}) = \max(1/4,\ |\lambda^2 + \epsilon^2|/2) \leq \max(1/4, (1 + \epsilon^2)/2) < 1$ *for* $\lambda \in \Gamma$. *This shows that* $\max_{\lambda \in \Gamma} r_\sigma(V(\lambda)T(\lambda)^{-1}) < 1$. *Hence by Theorem 3.5.3,* $T(\lambda) + V(\lambda)$ *has at most two eigenvalues (counting multiplicity) inside the circle* Γ. *Indeed, we have* $\tau := \sigma(T + V) \cap \mathrm{Int}(\Gamma) = \{\epsilon, -\epsilon\}$, *which consists of two simple eigenvalues of* $T(\lambda) + V(\lambda)$. *Hence* $m(0, T) = m(\tau, T + V) = 2$, *which validates Theorem 3.5.3. By contrast, for* $\lambda \in \Gamma$, *we have*

$$\|V(\lambda)T(\lambda)^{-1}\|_\infty = \max\left((2 + |41\lambda^3 + 3|)/8,\ |\lambda^2 + \epsilon^2|/2\right) \geq 5$$

which shows that $\max_{\lambda \in \Gamma} \|V(\lambda)T(\lambda)^{-1}\|_\infty \geq 5 > 1$. *Hence Theorem 3.3.8 cannot be invoked to conclude that* $T(\lambda) + V(\lambda)$ *has two eigenvalues (counting multiplicity) inside the circle* Γ. ■

For $V \in \mathbb{H}(\Omega, L(X))$, define $\|V\|_{\partial\Omega} := \frac{1}{2\pi}\int_{\partial\Omega} \|V(z)\|\,|dz|$. Since $z \longmapsto \|V(z)\|$ is subharmonic on Ω, it follows that $\|V\|_{\partial\Omega} = 0 \Longrightarrow V(z) = 0$ for all $z \in \Omega \cup \partial\Omega$. In fact, $\|\cdot\|_{\partial\Omega}$ defines a norm on $\mathbb{H}(\Omega, L(X))$. Next, define $\mathbb{V} : C(\partial\Omega, X) \longrightarrow C(\partial\Omega, X)$ by

$$\mathbb{V}f := \frac{1}{2\pi i}\int_\Gamma \mathcal{I}V(z)f(z)dz,$$

where $\mathcal{I}$ is the imbedding of X into $C(\partial\Omega, X)$. Then $\|\mathbb{V}f\|_\infty \leq \|V\|_{\partial\Omega}\|f\|_\infty$. Consider $W(\lambda, t) := T(\lambda)+tV(\lambda)$ for $t \in \mathbb{C}$. Then the linearization $\mathbb{T}(t)$ of $W(\lambda, t)$ can be written as $\mathbb{T}(t) = \mathbb{T} + t\mathbb{V}$ for $t \in \mathbb{C}$, where $\mathbb{T}$ is the linearization of $T(\lambda)$.

Let μ be a discrete eigenvalue of $T(\lambda)$ of algebraic multiplicity ℓ. Let $\Gamma \subset \rho(T)$ be such that $\sigma(T)\cap\mathrm{Int}(\Gamma) = \{\mu\}$. Let $\eth_\Gamma := \{t \in \mathbb{C} : \max_{z\in\Gamma} r_\sigma\left(V(z)T(z)^{-1}\right) < 1\}$. Then by Theorem 3.5.2, $W(\lambda, t)$ has ℓ discrete eigenvalues (counting multiplicity) $\mu_1(t), \ldots, \mu_\ell(t)$ inside the curve Γ. Set $\mu_{\mathrm{av}}(t) := (\mu_1(t) + \cdots + \mu_\ell(t))/\ell$. Then we have (see, [33, p. 405]) $\mu_{\mathrm{av}}(t) = \mu + \frac{1}{\ell}\mathrm{Tr}(\mathbb{VP})t + \mathcal{O}(|t|^2)$, where $\mathbb{P}$ is the spectral projection associated with $\mathbb{T}$ and λ. Hence we have the first order bound

$$|\mu_{\mathrm{av}}(t) - \mu| \leq \frac{1}{\ell}|\mathrm{Tr}(\mathbb{VP})|\,|t| + \mathcal{O}(|t|^2).$$

The one parameter family of operators $\mathbb{T}(t) = \mathbb{T}+t\mathbb{V}, t \in \mathbb{C}$, which is a linearization of the one parameter operator-valued function $W(\lambda, t)$, can be gainfully utilized to derive various perturbation bounds for discrete eigenvalues of $W(\lambda, t)$. Indeed, we have the following result.

Theorem 3.5.5. *Let μ be a discrete eigenvalue of $T(\lambda)$ of algebraic multiplicity ℓ and ascent ν. Let $\mu_1(t), \ldots, \mu_\ell(t)$ be the ℓ eigenvalues (counting multiplicities) of $W(\lambda, t)$ and $\mu_{\mathrm{av}}(t) := (\mu_1(t) + \cdots + \mu_\ell(t))/\ell$ for $t \in \eth_\Gamma$. Then the following hold.*

(a) There is a $\delta > 0$ and a constant C (independent of t and V) such that

$$\|tV\|_{\partial\Omega} < \delta \Longrightarrow |\mu_j(t) - \mu|^\nu \leq C\|tV\|_{\partial\Omega} \quad \textit{for} \quad j = 1, 2, \ldots, \ell.$$

(b) There is a $\delta > 0$ and a constant C (independent of t and V) such that

$$\|tV\|_{\partial\Omega} < \delta \Longrightarrow |\mu_{\mathrm{av}}(t) - \mu| \leq C\|tV\|_{\partial\Omega}.$$

Proof. Consider the operator $\mathbb{T}(t) := \mathbb{T} + t\mathbb{V}$ for $t \in \mathbb{C}$. Then it is well known that (see [39, 47, 1]) that there is a $\delta > 0$ and a constant C independent of t and $\mathbb{V}$ such that $\|t\mathbb{V}\|_\infty < \delta \Longrightarrow |\mu_j(t) - \mu|^\nu \leq C\|t\mathbb{V}\|_\infty$ for $j = 1 : \ell$. Similarly, there is a $\delta > 0$ and a constant C independent of t and $\mathbb{V}$ such that $\|t\mathbb{V}\|_\infty < \delta \Longrightarrow |\mu_{\mathrm{av}}(t) - \mu| \leq C\|t\mathbb{V}\|_\infty$. Now the desired results follow from the fact that $\|\mathbb{V}\|_\infty \leq \|V\|_{\partial\Omega}$. □

Let X^* denote the conjugate dual space of a complex Banach space X. For $x \in X$ and $x^* \in X^*$, we define $\langle x, x^* \rangle := x^*(x)$. Let $T \in L(X)$. The conjugate dual operator $T^* \in L(X^*)$ of T is given by

$$\langle Tx, x^* \rangle = \langle x, T^*x^* \rangle \text{ for all } x \in X \text{ and } x^* \in X^*.$$

It is well known that $\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}$ and λ is a discrete eigenvalue of $T \iff \bar{\lambda}$ is a discrete eigenvalue of T^*. Further, the algebraic (resp., geometric) multiplicity of λ as an eigenvalue of T is the same as the algebraic (resp., geometric) multiplicity of $\bar{\lambda}$ as an eigenvalue of T^*; see [39, p.112-115].

Definition 3.5.6. *Let $T \in L(X)$. Then $(\lambda, \phi, \phi^*) \in \mathbb{C} \times X \times X^*$ with ϕ and ϕ^* being nonzero is said to be a simple eigentriple of T if $T\phi = \lambda\phi$ and $T^*\phi^* = \bar{\lambda}\phi^*$ and λ is a simple eigenvalue of T.*

Let λ be a discrete eigenvalue of T. Let P be the spectral projection of T corresponding to the eigenvalue λ and let S be the reduced resolvent of $T - \lambda I$. Then we have the following bound when λ is simple.

Theorem 3.5.7 ([46]). *Let $T, V \in L(X)$. Let (λ, ϕ, ϕ^*) be a simple eigentriple of T such that $\langle \phi, \phi^* \rangle = 1$. Then for every real number $r > 0$ satisfying*

$$\beta_V := \max\{\|PV\|, \|(I - P)V\|\} \leq \frac{r}{\|S\|(1 + r)^2},$$

the perturbed operator $T + V$ has a simple eigenvalue μ and a corresponding (unique) eigenvector ψ such that $\langle \psi, \phi^* \rangle = 1$,

$$\|\psi - \phi\| \leq r \quad \textit{and} \quad |\mu - \lambda| \leq \|P\|(1 + \|\psi - \phi\|)\|V\|.$$

Further, λ is the only eigenvalue of $T + V$ lying inside the disc

$$\Delta := \{z \in \mathbb{C} : \ |z - q| < \frac{\mathrm{d}}{2}(1 + \sqrt{1 - 4\alpha})\},$$

where $\mathrm{d} := \dfrac{(1 - 2\beta_V\|S\|)}{\|S\|}$, $q := \lambda + \langle V\psi, \phi^* \rangle$ *and* $\alpha := \left(\dfrac{r}{1 + r^2}\right)^2$.

We now derive a similar bound for a simple eigenvalue of $T \in \mathbb{H}(\Omega, L(X))$. Let μ be a discrete eigenvalue of $T(\lambda)$ and $\mathbb{T}$ be a linearization of $T(\lambda)$ as given in Theorem 3.4.3. Let $\mathbb{P}$ be the spectral projection of $\mathbb{T}$ corresponding to μ and $\mathbb{S}$ be the reduced resolvent of $\mathbb{T} - \mu I$. Set $\mathbf{s} := \|\mathbb{S}\|_\infty$ and $\mathbf{p} := \|\mathbb{P}\|_\infty$. We denote by $\mathbb{T}^*$ the conjugate dual of the operator $\mathbb{T}$. Then we have the following bound which follows from Theorem 3.5.7.

Theorem 3.5.8. *Let μ be a simple eigenvalue of $T(\lambda)$. Let $\mathbb{T}\Phi = \mu\Phi$ and $\mathbb{T}^*\Phi^* = \bar{\mu}\Phi^*$ be such that $\langle \Phi, \Phi^* \rangle = 1$. Then for every real number $r > 0$ and $V \in \mathbb{H}(\Omega, L(X))$ satisfying*

$$\beta_V := \max\{\|\mathbb{P}\mathbb{V}\|_\infty, \|(I - \mathbb{P})\mathbb{V}\|_\infty\} \leq \|V\|_{\partial\Omega} \leq \frac{r}{\mathbf{s}(1+r)^2},$$

the operator-valued function $W(\lambda) := T(\lambda) + V(\lambda)$ has a simple eigenvalue μ_V and a corresponding (unique) eigenvector Ψ such that $\langle \Psi, \Phi^* \rangle = 1$,

$$\|\Psi - \Phi\|_\infty \leq r \quad \textit{and} \quad |\mu_V - \mu| \leq \mathbf{p}(1 + \|\Psi - \Phi\|_\infty)\|V\|_{\partial\Omega}.$$

Further, μ_V is the only eigenvalue of $W(\lambda)$ lying inside the disc

$$\Delta := \{z \in \mathbb{C} : \ |z - q| < \frac{\mathbf{d}}{2}(1 + \sqrt{1 - 4\alpha})\},$$

where $\mathbf{d} := \dfrac{(1 - 2\beta_V \mathbf{s})}{\mathbf{s}}$, $q := \mu + \langle \mathbb{V}\Phi, \Phi^* \rangle$ *and* $\alpha := \left(\dfrac{r}{1+r^2}\right)^2$.

The next result gives an upper bound of $|t|$ that ensures $t \in \partial_\Gamma$ which in turn yields bounds for perturbed eigenvalues and eigenvectors.

Theorem 3.5.9. *Let (μ, v) be a simple eigenpair of $T(\lambda)$. Let $\mathbf{p}$ and $\mathbf{s}$ be as in Theorem 3.5.8. Set $\beta := \max\{\|\mathbb{P}\|_\infty, \ \|I - \mathbb{P}\|_\infty\}$. Let $V \in \mathbb{H}(\Omega, L(X))$. Then for every real number $r \geq 0$ and $t \in \mathbb{C}$ satisfying*

$$|t| \leq \frac{2\pi r}{(1+r)^2 \beta\, \mathbf{s}\, \|V\|_{\partial\Omega}}$$

the operator-valued function $T(z) + tV(z)$ has a simple eigenvalue $\lambda(t)$ and a unique eigenvector $v(t)$ such that

$$\|\mathcal{E}_{\lambda(t)} v(t) - \mathcal{E}_\mu v\| \leq r \ \textit{and} \ |\lambda(t) - \mu| \leq \mathbf{p}|t|(1+r)\|V\|_{\partial\Omega},$$

where the maps $\mathcal{E}_{\lambda(t)} : \mathcal{N}(T(\lambda(t)) + tV(\lambda(t))) \longrightarrow \mathcal{N}(\mathbb{T}(t) - \lambda(t)I), \ u \longmapsto \mathcal{E}_{\lambda(t)}u$, and $\mathcal{E}_\mu : \mathcal{N}(T(\mu)) \longrightarrow \mathcal{N}(\mathbb{T} - \mu I), \ v \longmapsto \mathcal{E}_\mu v$, are isomorphisms given by

$$(\mathcal{E}_{\lambda(t)}u)(w) := \frac{u}{\lambda(t) - w} \ \textit{and} \ (\mathcal{E}_\mu v)(w) := \frac{v}{\mu - w}, \ w \in \partial\Omega.$$

Proof. Let $\mathbb{T}$ be the linearization of $T(\lambda)$. Set $\Phi := \mathcal{E}_\mu v$. Recall that $\mathcal{E}_\mu$ is defined in Proposition 3.4.5. Then (μ, Φ) is a simple eigenpair of $\mathbb{T}$. Let $\mathbb{T}(t)$ be the linearization of

$T(\lambda)+tV(\lambda)$ for $t \in \mathbb{C}$. Set $\mathbb{V}(t) := \mathbb{T}(t)-\mathbb{T}$ for $t \in \mathbb{C}$. Then for $|t| \leq \dfrac{2\pi r}{(1+r)^2\beta_V \mathbf{s}\|V\|_{\partial\Omega}}$,

$$\begin{aligned}\|\mathbb{V}(t)\|_\infty = \|\mathbb{T}(t)-\mathbb{T}\|_\infty &= \sup_{\|f\|_\infty\leq 1} \|(\mathbb{T}(t)-\mathbb{T})f\|_\infty, \; f \in C(\partial\Omega)^n \\ &= \sup_{\|f\|_\infty\leq 1} \|\frac{1}{2\pi}\int_{\partial\Omega}(-tV(z))f(z)dz\| \\ &\leq \frac{|t|}{2\pi}\|V\|_{\partial\Omega} \leq \frac{r}{\beta_V \mathbf{s}(1+r)^2}.\end{aligned}$$

This shows that

$$\max\{\|\mathbb{P}\mathbb{V}(t)\|_\infty, \; \|(I-\mathbb{P})\mathbb{V}(t)\|_\infty\} \leq \|\mathbb{V}(t)\|_\infty \beta_V \leq \frac{r}{\mathbf{s}(1+r)^2}.$$

By Theorem 3.5.8, there is a unique simple eigenpair $(\lambda(t), \Phi(t))$ of $\mathbb{T}(t) = \mathbb{T} + \mathbb{V}(t)$ such that

$$\|\Phi(t)-\Phi\|_\infty \leq r \text{ and } |\lambda(t)-\mu| \leq \mathbf{p}(1+\|\Phi(t)-\Phi\|_\infty)\|tV\|_{\partial\Omega}.$$

Now, the desired bounds follow from the fact that there exists a unique eigenvector $v(t)$ of $T(\lambda)+tV(\lambda)$ corresponding to $\lambda(t)$ such that $\Phi(t) = \mathcal{E}_{\lambda(t)}v(t)$.

□

4 Realization and Linearization of Meromorphic Matrix-valued Functions

Let $G(z)$ be an $m \times n$ strictly proper rational matrix, that is, $G(z) \to 0$ as $z \to \infty$. Then $G(z)$ admits a transfer function realization [15, 32] of the form

$$G(z) = C(zI - A)^{-1}B$$

for some matrices A, B and C. The matrix triple (C, A, B) is called a realization of $G(z)$. The realization (C, A, B) is called minimal when the size of A is the smallest. In such a case, $\wp_{\mathbb{C}}(G) = \mathrm{eig}(A)$. An $m \times n$ rational matrix $G(z)$ admits a transfer function realization [32, 50, 56] of the form

$$G(z) = P(z) + C(zI - A)^{-1}B$$

for some $P(z) \in \mathbb{C}[z]^{m \times n}$. In such a case, zeros and poles of $G(z)$ can be analyzed via a linearization [2, 3, 14, 4] of $G(z)$ of the form

$$\mathbb{L}(z) := \left[\begin{array}{c|c} \mathcal{X} - z\mathcal{Y} & \mathcal{C} \\ \hline \mathcal{B} & A - zI \end{array}\right],$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{B}, \mathcal{C}$ are appropriate matrices. The main aim of this chapter is to generalize the template

rational matrix $\longrightarrow$ realization $\longrightarrow$ linearization $\longrightarrow$ spectral analysis

to the case of meromorphic matrix-valued functions.

Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m \times n}$ and $\mathcal{O} \subset \Omega$ be a compact region. We prove that $\mathbf{M}(z)$ admits a local realization of the form

$$\mathbf{M}(z) = \mathbf{H}(z) + C(zI - A)^{-1}B$$

for some $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ and show that (C, A, B) can be computed from the Markov parameters associated with $\mathbf{M}(z)$. In the special case when $\mathbf{M}(z)$ is regular, we construct a linearization of $\mathbf{M}(z)$ of the form

$$\mathbb{L}(z) := \left[\begin{array}{c|c} zI - \mathbb{T} & \mathcal{C} \\ \hline \mathcal{B} & A - zI \end{array}\right]$$

and show that $\sigma_\Omega(\mathbf{M}) \cap \mathcal{O} = \sigma(\mathbb{T}) \cap \mathcal{O}$, where $\mathcal{C}, \mathbb{T}$ and $\mathcal{B}$ are appropriate bounded linear operators on some Banach spaces.

4.1 Local realization of meromorphic matrices

Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\lambda \in \wp_\Omega(\mathbf{M})$. Then there exists $\nu \in \mathbb{N}$ such that the Laurent series of $\mathbf{M}(z)$ at λ is given by

$$\mathbf{M}(z) = \sum_{j=0}^{\infty} A_j (z-\lambda)^j + \sum_{j=1}^{\nu} \frac{D_j}{(z-\lambda)^j}, \tag{4.1}$$

where $D_\nu \neq 0$. Note that ν is the order of the pole λ and $G_\lambda(z) := \sum_{j=1}^{\nu} \frac{D_j}{(z-\lambda)^j}$ is the principal part of the Laurent series of $\mathbf{M}(z)$ at λ. Also note that $G_\lambda(z)$ is an $m \times n$ rational matrix which is strictly proper, that is, $G_\lambda(z) \longrightarrow 0$ as $z \to \infty$. Now consider the polynomial $P(z) := \sum_{j=1}^{\nu} D_j z^j \in \mathbb{C}[z]^{m\times n}$. Then $G_\lambda(z) = P(\frac{1}{z-\lambda})$. This motivates us to define the principal pole polynomial of a meromorphic matrix corresponding to a pole.

Definition 4.1.1. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ *and* $\lambda \in \wp_\Omega(\mathbf{M})$. *Then a matrix polynomial* $P(z) \in \mathbb{C}[z]^{m\times n}$ *is said to be the principal pole polynomial of* $\mathbf{M}(z)$ *at* λ *if* $P(0) = 0$ *and* $M(z) - P\left(\frac{1}{z-\lambda}\right)$ *is analytic at* λ.

We have the following result which will be useful later.

Proposition 4.1.2. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ *and* $\lambda \in \wp_\Omega(\mathbf{M})$. *Then* $P(z)$ *is the principal pole polynomial of* $\mathbf{M}(z)$ *at* $\lambda \Longleftrightarrow P(\frac{1}{z-\lambda})$ *is the principal part of the Laurent series of* $\mathbf{M}(z)$ *at* λ. *Further, if* $P(z)$ *is the principal pole polynomial* $\mathbf{M}(z)$ *at* λ *then* $\deg(P(z))$ *is the order of the pole* λ *and*

$$\lim_{z\to\lambda} \left[\mathbf{M}(z) - P\left(\frac{1}{z-\lambda}\right)\right] = \frac{1}{2\pi i}\int_\Gamma \frac{\mathbf{M}(z)}{z-\lambda} dz,$$

where Γ *is a positively oriented rectifiable simple closed curve in* Ω *isolating* λ *from the rest of the poles of* $\mathbf{M}(z)$.

Proof. If $P(\frac{1}{z-\lambda})$ is the principal part of Laurent series of $\mathbf{M}(z)$ at λ then, in view of (4.1), $\mathbf{M}(z) - P(\frac{1}{z-\lambda})$ is analytic at λ and $P(0) = 0$. Hence $P(z)$ is the principal pole polynomial of $\mathbf{M}(z)$ at λ. Conversely, if $P(z)$ is the principal pole polynomial of $\mathbf{M}(z)$ at λ then $\mathbf{M}(z) - P(\frac{1}{z-\lambda})$ is analytic at λ. By Taylor series expansion at λ, we have

$$\mathbf{M}(z) - P(\frac{1}{z-\lambda}) = \sum_{j=0}^{\infty} A_j(z-\lambda)^j \Longrightarrow \mathbf{M}(z) = \sum_{j=0}^{\infty} A_j(z-\lambda)^j + P(\frac{1}{z-\lambda})$$

is the Laurent series of $\mathbf{M}(z)$ at λ. Hence $P(\frac{1}{z-\lambda})$ is the principal part of the Laurent series of $\mathbf{M}(z)$ at λ. Evidently, $\deg(P(z))$ is the order of the pole λ.

Now, consider the Laurent series $\mathbf{M}(z) = \sum_{j=0}^{\infty} A_j(z-\lambda)^j + P(\frac{1}{z-\lambda})$ of $\mathbf{M}(z)$ at λ. Then by Cauchy's integral formula, we have $\int_\Gamma \frac{P(\frac{1}{z-\lambda})}{z-\lambda} dz = 0$ and hence

$$\frac{1}{2\pi i}\int_\Gamma \frac{\mathbf{M}(z)}{z-\lambda} dz = A_0 = \lim_{z\to\lambda}\left[\mathbf{M}(z) - P\left(\frac{1}{z-\lambda}\right)\right].$$

□

Let $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathbf{H} \in \mathbb{H}(\Omega)^{m\times n}$. If $\mathbf{M}_1(z) = \mathbf{M}_2(z) + \mathbf{H}(z)$ then $\wp_\Omega(\mathbf{M}_1) = \wp_\Omega(\mathbf{M}_2)$ and the Laurent series of $\mathbf{M}_1(z)$ and $\mathbf{M}_2(z)$ at each pole have the same principal part. Thus, the pole structure of a meromorphic matrix is invariant under additive analytic perturbation. Hence, for analysis of poles, we define an equivalence relation that preserves pole structures of meromorphic matrices.

Definition 4.1.3. *Let $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathcal{O} \subset \Omega$ be open. Then $\mathbf{M}_1$ and $\mathbf{M}_2$ are said to be* **pole equivalent** *on $\mathcal{O}$ and written as $\mathbf{M}_1(z) \simeq_{\mathcal{O}} \mathbf{M}_2(z)$ if $\mathbf{M}_1(z) - \mathbf{M}_2(z)$ is analytic in $\mathcal{O}$, that is, if $\mathbf{M}_1 - \mathbf{M}_2 \in \mathbb{H}(\mathcal{O})^{m\times n}$.*

Let $\lambda \in \Omega$. Then $\mathbf{M}_1$ and $\mathbf{M}_2$ are said to be **pole equivalent** *at λ and written as $\mathbf{M}_1(z) \simeq_\lambda \mathbf{M}_2(z)$ if $\mathbf{M}_1(z) - \mathbf{M}_2(z)$ is analytic at λ, that is, if $\mathbf{M}_1 - \mathbf{M}_2 \in \mathbb{H}(\mathcal{O})^{m\times n}$ for some open set $\mathcal{O} \subset \Omega$ containing λ.*

Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\lambda \in \wp_\Omega(\mathbf{M})$. Let $P(z) \in \mathbb{C}[z]^{m\times n}$ be the principal pole polynomial of $\mathbf{M}(z)$ at λ. Then by Proposition 4.1.2, we have $\mathbf{M}(z) \simeq_\lambda P\left(\frac{1}{z-\lambda}\right)$. Thus, as far as poles are concerned, $\mathbf{M}(z)$ is locally represented in a neighbourhood of each pole by a strictly proper rational matrix. More generally, we have the following result.

Theorem 4.1.4. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathcal{O} \subset \Omega$ be open. Suppose that $\wp_{\mathcal{O}}(\mathbf{M}) = \wp_\Omega(\mathbf{M}) \cap \mathcal{O} = \{\lambda_1, \ldots, \lambda_\ell\}$. Let $P_1, \ldots, P_\ell$ be the principal pole polynomials of $\mathbf{M}(z)$ at*

$\lambda_1, \ldots, \lambda_\ell$, respectively. Then there exists $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ such that

$$\mathbf{M}(z) = \mathbf{H}(z) + P_1\left(\frac{1}{z-\lambda_1}\right) + \cdots + P_\ell\left(\frac{1}{z-\lambda_\ell}\right) \quad \text{for } z \in \mathcal{O},$$

that is, $\mathbf{M}(z) \simeq_{\mathcal{O}} \left[P_1\left(\frac{1}{z-\lambda_1}\right) + \cdots + P_\ell\left(\frac{1}{z-\lambda_\ell}\right)\right]$.

Set $G(z) := P_1\left(\frac{1}{z-\lambda_1}\right) + \cdots + P_\ell\left(\frac{1}{z-\lambda_\ell}\right)$. Then $G(z)$ is the unique strictly proper rational matrix such that $\mathbf{M}(z) \simeq_{\mathcal{O}} G(z)$ and $\wp_{\mathcal{O}}(\mathbf{M}) = \wp_{\mathbb{C}}(G)$.

Proof. Define $\mathbf{H}(z) := \mathbf{M}(z) - \left[P_1\left(\frac{1}{z-\lambda_1}\right) + \cdots + P_\ell\left(\frac{1}{z-\lambda_\ell}\right)\right]$ for $z \in \mathcal{O}$. By Proposition 4.1.2, we have $\mathbf{M}(z) \simeq_{\lambda_j} P_j\left(\frac{1}{z-\lambda_j}\right)$ for $j = 1 : \ell$. Hence $\mathbf{H}(z)$ has removable singularities at $\lambda_1, \ldots, \lambda_\ell$. By analytic continuation, we conclude that $\mathbf{H}(z)$ is analytic in $\mathcal{O}$. Indeed, setting

$$\mathbf{H}(\lambda_j) := \lim_{z\to\lambda_j}\left[\mathbf{M}(z) - \sum_{k=1}^{\ell} P\left(\frac{1}{z-\lambda_k}\right)\right] = \frac{1}{2\pi i}\int_{\Gamma_j}\frac{\mathbf{M}(z)}{z-\lambda_j}dz - \sum_{k=1,k\neq j}^{\ell} P\left(\frac{1}{\lambda_j-\lambda_k}\right),$$

we have $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$. Here Γ_j is a simple closed rectifiable positively oriented curve in $\mathcal{O}$ isolating λ_j from the rest of the poles of $\mathbf{M}(z)$. This proves that $\mathbf{M}(z) \simeq_{\mathcal{O}} G(z)$.

By construction, $G(z)$ is a unique strictly proper rational matrix such that

$$\mathbf{M}(z) \simeq_{\mathcal{O}} G(z) \quad \text{and} \quad \wp_{\mathcal{O}}(\mathbf{M}) = \wp_{\mathbb{C}}(G).$$

□

Definition 4.1.5. *[15] Let $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$. The pair (C, A) is said to be observable if the observability matrix* $\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1}\end{bmatrix}$ *has full column rank.*

The pair (A, B) is said to be controllable if the controllablity matrix $\begin{bmatrix} A & AB & \cdots & A^{p-1}B\end{bmatrix}$ has full row rank.

A useful characterization of observable/controllable pair is as follows.

Theorem 4.1.6. *[15] Let $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$. Then*

$$(C, A) \text{ is observable} \iff \operatorname{rank}\left(\begin{bmatrix} A - zI \\ C\end{bmatrix}\right) = p \text{ for all } z \in \mathbb{C},$$

$$(A, B) \text{ is controllable} \iff \operatorname{rank}\left(\begin{bmatrix} A - zI & B\end{bmatrix}\right) = p \text{ for all } z \in \mathbb{C}.$$

Realization theory of rational matrices is a classical subject and has been studied extensively over the years, see [32, 15, 56, 50] and the references therein. A realization of a strictly proper rational matrix is defined as follows.

Definition 4.1.7. *[15] Let $G(z) \in \mathbb{C}(z)^{m\times n}$ be a strictly proper rational matrix. Then $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$ is said to be a realization of $G(z)$ if*

$$G(z) = C(zI_p - A)^{-1}B \quad \text{for} \quad z \in \mathbb{C}.$$

The realization (C, A, B) is said to be minimal if (C, A) is observable and (A, B) is controllable.

For a ready reference, we collect some well-known results in the following theorem which will be useful for local analysis of meromorphic matrices.

Theorem 4.1.8. *[32, 15] Let $G(z) \in \mathbb{C}(z)^{m\times n}$ be a strictly proper rational matrix. Then there exists a minimal realization $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$, for some $p \in \mathbb{N}$, such that*

$$G(z) = C(zI_p - A)^{-1}B \quad \text{and} \quad \wp_{\mathbb{C}}(G) = \mathrm{eig}(A).$$

The Rosenbrock system matrix

$$\mathbf{S}(z) := \left[\begin{array}{c|c} 0 & C \\ \hline B & A - z_pI \end{array}\right]$$

is irreducible (minimal order) and $\sigma_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(\mathbf{S})$.

If $S \in \mathbb{C}^{p\times p}$ is nonsingular then $(CS,\ S^{-1}AS,\ S^{-1}B)$ is a minimal realization of $G(z)$. If (C_1, A_1, B_1) and (C_2, A_2, B_2) are minimal realizations of $G(z)$ then exists a unique nonsingular matrix S such that $C_2 = C_1S, A_2 = S^{-1}A_1S$ and $B_2 = S^{-1}B_1$.

Remark 4.1.9. *If $G(z) \in \mathbb{C}(z)^{m\times n}$ is non-proper then there exists a matrix polynomial $P(z) \in \mathbb{C}[z]^{m\times n}$ such that $G(z)-P(z)$ is a strictly proper rational matrix [32, 50]. Hence there exists a minimal realization $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$, for some $p \in \mathbb{N}$, such that*

$$G(z) = P(z) + C(zI_p - A)^{-1}B \quad \text{and} \quad \wp_{\mathbb{C}}(G) = \mathrm{eig}(A).$$

Further, the Rosenbrock system matrix

$$\mathbf{S}(z) := \left[\begin{array}{c|c} P(z) & C \\ \hline B & A - zI_p \end{array}\right]$$

is irreducible (minimal order) and $\sigma_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(\mathbf{S})$.

For local analysis, we now define local realizations of meromorphic matrices.

Definition 4.1.10. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ *and* $\mathcal{O} \subset \Omega$ *be open.*

(a) Then $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$*, for some* $p \in \mathbb{N}$*, is said to be a local realization of* $\mathbf{M}(z)$ *on* $\mathcal{O}$ *if*

$$\mathbf{M}(z) \simeq_{\mathcal{O}} C(zI_p - A)^{-1}B \text{ and } \sigma(A) \subset \mathcal{O}.$$

If (C, A) *is observable and* (A, B) *is controllable then* (C, A, B) *is called a minimal local realization (MLR) of* $\mathbf{M}(z)$ *on* $\mathcal{O}$*.*

Let $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ *be such that* $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_p - A)^{-1}B$ *for* $z \in \mathcal{O}$*. The matrix*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_p \end{array}\right] \text{ for } z \in \mathcal{O}$$

is called the system matrix in state-space form (SSF) of $\mathbf{M}(z)$ *on* $\mathcal{O}$*. The system matrix* $\mathbf{S}(z)$ *is called irreducible when* (C, A, B) *is a minimal realization of* $\mathbf{M}(z)$ *on* $\mathcal{O}$*.*

(b) Let $\lambda \in \wp_\Omega(\mathbf{M})$*. Then* $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$*, for some* $p \in \mathbb{N}$*, is said to be a local realization of* $\mathbf{M}(z)$ *at* λ *if*

$$\mathbf{M}(z) \simeq_{\lambda} C(zI_p - A)^{-1}B \text{ and } \sigma(A) = \{\lambda\}.$$

If (C, A) *is observable and* (A, B) *is controllable then* (C, A, B) *is called a minimal local realization of* $\mathbf{M}(z)$ *at* λ*.*

If (C, A, B) is a minimal local realization of $\mathbf{M}(z)$ at λ then there exists an open set $\mathcal{O} \subset \Omega$ and $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ such that $\mathbf{M}(z) = \mathbf{H}(z) + C(zI - A)^{-1}B$ for $z \in \mathcal{O}$.

Remark 4.1.11. *(a) A local realization* $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$ *of* $\mathbf{M}(z)$ *on* $\mathcal{O}$ *is called a standard local realization. A nonstandard local realization of* $\mathbf{M}(z)$ *on* $\mathcal{O}$ *is of the form*

$$\mathbf{M}(z) = \mathbf{H}(z) + C(A - zE)^{-1}B \text{ and } \mathrm{eig}(A, E) \subset \mathcal{O},$$

where $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ *and* $E \in \mathbb{C}^{p\times p}$ *is nonsingular. Here* $\mathrm{eig}(A, E)$ *is the spectrum of the pencil* $A - zE$*. In such a case, the system matrix is given by*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & zE - A \end{array}\right] \text{ for } z \in \mathcal{O}.$$

A nonstandard realization is said to be minimal when

$$\operatorname{rank}\left(\begin{bmatrix} A - zE & B \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} A - zE \\ C \end{bmatrix}\right) = p$$

for all $z \in \mathbb{C}$. In such a case the system matrix $\mathbf{S}(z)$ is called irreducible.

We have the following result whose proof is immediate; see Theorem 4.1.8.

Theorem 4.1.12. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathcal{O} \subset \Omega$ be open. Then we have the following.*

(a) *If $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$ is a (minimal) local realization of $\mathbf{M}(z)$ on $\mathcal{O}$ then so is $(CS,\ S^{-1}AS,\ S^{-1}B)$ for any nonsingular matrix $S \in \mathbb{C}^{p\times p}$.*

(b) *If $\mathbf{M}(z) \simeq_{\mathcal{O}} C(zE - A)^{-1}B$ is a nonstandard local realization of $\mathbf{M}(z)$ then both $(C, E^{-1}A,\ E^{-1}B)$ and $(CE^{-1},\ AE^{-1},\ B)$ are standard local realizations of $\mathbf{M}(z)$ on $\mathcal{O}$.*

(c) *Let (C_1, A_1, B_1) and (C_2, A_2, B_2) be minimal local realizations of $\mathbf{M}(z)$ on $\mathcal{O}$. Then there exists a unique nonsingular matrix S such that $C_2 = C_1S, A_2 = S^{-1}A_1S$ and $B_2 = S^{-1}B_1$.*

Let $\mathcal{O} \subset \Omega$ be open. Suppose that $\wp_{\mathcal{O}}(\mathbf{M}) = \{\lambda_1, \ldots, \lambda_\ell\}$. It is easy to see that if $\mathbf{M}(z) \simeq_{\lambda_j} C_j(zI - A_j)^{-1}B_j$ is a local realization at λ_j for $j = 1 : \ell$ then

$$\mathbf{M}(z) \simeq_{\mathcal{O}} \left[C_1(zI - A_1)^{-1}B_1 + \cdots + C_\ell(zI - A_\ell)^{-1}B_\ell\right].$$

Here the generic notation I is the identity matrix having the same size as that of A_j. We have the following result.

Theorem 4.1.13. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathcal{O} \subset \Omega$ be open. Suppose that $\wp_{\mathcal{O}}(\mathbf{M}) = \{\lambda_1, \ldots, \lambda_\ell\}$. Let $(C_1, A_1, B_1), \cdots, (C_\ell, A_\ell, B_\ell)$ be minimal local realization of $\mathbf{M}(z)$ at $\lambda_1, \ldots, \lambda_\ell$, respectively. Set*

$$C := \begin{bmatrix} C_1 & \cdots & C_\ell \end{bmatrix}, A := \operatorname{diag}(A_1, \ldots, A_\ell) \text{ and } B := \begin{bmatrix} B_1 \\ \vdots \\ B_\ell \end{bmatrix}.$$

Then (C, A, B) is a minimal local realization of $\mathbf{M}(z)$ on $\mathcal{O}$, that is,

$$\mathbf{M}(z) \simeq_{\mathcal{O}} \left[C_1(zI - A_1)^{-1}B_1 + \cdots + C_\ell(zI - A_\ell)^{-1}B_\ell\right] = C(zI - A)^{-1}B$$

and (C, A) is observable and (A, B) is controllable.

Proof. Since $\mathbf{M}(z) - \sum_{j=1}^{\ell} C_j(zI - A_j)^{-1}B_j$ has removable singularities at $\lambda_1, \ldots, \lambda_\ell$, by considering its analytic continuation, we have

$$\mathbf{M}(z) \simeq_{\mathcal{O}} \left[C_1(zI - A_1)^{-1}B_1 + \cdots + C_\ell(zI - A_\ell)^{-1}B_\ell\right] = C(zI - A)^{-1}B.$$

Since $(C_1, A_1), \ldots, (C_\ell, A_\ell)$ are observable, it is easily seen that (C, A) is observable. Indeed, it is enough to show that $\begin{bmatrix} \lambda_j I - A \\ C \end{bmatrix}$ has full column rank for $j = 1 : \ell$. Since $\lambda_j I - A_i$ is nonsingular for $i \neq j$ and $\begin{bmatrix} \lambda_j I - A_j \\ C_j \end{bmatrix}$ has full column rank for $j = 1 : \ell$, it follows that

$$\begin{bmatrix} \lambda_j I - A \\ C \end{bmatrix} = \begin{bmatrix} \lambda_j I - A_1 & & & \\ & \lambda_j I - A_2 & & \\ & & \ddots & \\ & & & \lambda_j I - A_\ell \\ C_1 & C_2 & \cdots & C_\ell \end{bmatrix}$$

has full column rank. Hence (C, A) is observable.

Similarly, as $(A_1, B_1), \ldots, (A_\ell, B_\ell)$ are controllable, it is easily seen that (A, B) is controllable. Hence (C, A, B) is a minimal local realization of $\mathbf{M}(z)$ on $\mathcal{O}$. $\square$

We now show that the existence of a minimal local realization of $\mathbf{M}(z)$ follows from the principal pole polynomials of $\mathbf{M}(z)$. Let $\#(S)$ denote the cardinality of a set S. If $\#(S) < \infty$ then S is a finite set and $\#(S)$ is the number of elements of the set S.

Theorem 4.1.14. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ and $\mathcal{O} \subset \Omega$ be open. Suppose that $\#(\wp_{\mathcal{O}}(\mathbf{M})) < \infty$. Then there exists $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$, for some $p \in \mathbb{N}$, and $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ such that $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_p - A)^{-1}B$ is a minimal local realization of $\mathbf{M}(z)$ on $\mathcal{O}$ and $\wp_{\mathcal{O}}(\mathbf{M}) = \mathrm{eig}(A) = \wp_{\Omega}(\mathbf{M}) \cap \mathcal{O}$. Consider the system matrix in SSF*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_p \end{array}\right] \quad \text{for } z \in \mathcal{O}.$$

Then $S(z)$ is irreducible on $\mathcal{O}$ and $\sigma_{\mathcal{O}}(\mathbf{M}) = \sigma_{\mathcal{O}}(\mathbf{S}) = \sigma_{\Omega}(\mathbf{M}) \cap \mathcal{O}$.

Proof. Suppose that $\wp_{\mathcal{O}}(\mathbf{M}) = \{\lambda_1, \ldots, \lambda_\ell\}$. Let $P_1(z), \ldots, P_\ell(z)$ be principal pole polynomials of $\lambda_1, \ldots, \lambda_\ell$, respectively. Then by Theorem 4.1.4, there exists $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$

such that

$$\mathbf{M}(z) = \mathbf{H}(z) + P_1\left(\frac{1}{z-\lambda_1}\right) + \cdots + P_\ell\left(\frac{1}{z-\lambda_\ell}\right) \quad \text{for} \;\; z \in \mathcal{O}.$$

Further, $G(z) := P_1\left(\frac{1}{z-\lambda_1}\right)+\cdots+P_\ell\left(\frac{1}{z-\lambda_\ell}\right)$ is the unique strictly proper rational matrix such that $\mathbf{M}(z) \simeq_{\mathcal{O}} G(z)$ and $\wp_{\mathcal{O}}(\mathbf{M}) = \wp_{\mathbb{C}}(G)$. Since $G(z)$ is a strictly proper rational matrix, by Theorem 4.1.8, there exists $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$, for some $p \in \mathbb{N}$, such that $G(z) = C(zI_p - A)^{-1}B$ is a minimal realization and $\wp_{\mathbb{C}}(G) = \mathrm{eig}(A)$. This shows that $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_p - A)^{-1}B$ is a minimal local realization of $\mathbf{M}(z)$ on $\mathcal{O}$ and $\wp_{\mathcal{O}}(\mathbf{M}) = \wp_{\mathbb{C}}(G) = \mathrm{eig}(A)$.

Obviously, the system matrix $\mathbf{S}(z)$ is irreducible as (C, A) is observable and (A, B) is controllable. By Theorem 2.4.18, we have $\sigma_{\mathcal{O}}(\mathbf{M}) = \sigma_{\mathcal{O}}(\mathbf{S}) = \sigma_{\Omega}(\mathbf{M}) \cap \mathcal{O}$. □

Observe that $(C, A, B) \in \mathbb{C}^{m\times p} \times \mathbb{C}^{p\times p} \times \mathbb{C}^{p\times n}$, for some $p \in \mathbb{N}$, is a local realization of $\mathbf{M}(z)$ at a pole $\lambda \Longleftrightarrow C(zI_p - A)^{-1}B$ is the principal part of the Laurent series of $\mathbf{M}(z)$ at λ. The principal pole polynomial $\mathbf{M}(z)$ at a pole can be recovered from a minimal local realization of $\mathbf{M}(z)$. We have the following result whose proof is immediate.

Proposition 4.1.15. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ *and* $\mathcal{O} \subset \Omega$ *be open. Suppose that* $\wp_{\mathcal{O}}(\mathbf{M}) = \{\lambda_1, \ldots, \lambda_\ell\}$ *and* $\mathbf{M}(z) \simeq_{\mathcal{O}} C(zI - A)^{-1}B$ *is a minimal local realization of* $\mathbf{M}(z)$.

(a) There exists a nonsingular matrix S *such that* $S^{-1}AS = \mathrm{diag}(A_1, \ldots, A_\ell)$ *and* $\mathrm{eig}(A_j) = \{\lambda_j\}$ *for* $j = 1 : \ell$. *Let* $S = \begin{bmatrix} X_1 & \cdots & X_\ell \end{bmatrix}$ *and* $(S^{-1})^* = \begin{bmatrix} Y_1 & \cdots & Y_\ell \end{bmatrix}$ *be conformal partitions. Then* $\mathbf{M}(z) \simeq_{\lambda_j} CX_j(zI - A_j)^{-1}Y_j^*B$ *is a minimal local realization of* $\mathbf{M}(z)$ *at* λ_j *for* $j = 1 : \ell$.

(b) Let ν_j *be the order of the pole* λ_j *for* $j = 1 : \ell$. *Then* ν_j *the ascent of* λ_j *as an eigenvalue of* A *and*

$$P_j(z) := CX_j(z^{-1}I + \lambda_j I - A_j)^{-1}Y_j^*B = \sum_{k=1}^{\nu_j - 1} z^{k+1} CX_j(A - \lambda_j I)^k Y_j^* B$$

is the principal pole polynomial of $\mathbf{M}(z)$ *at* λ_j *for* $j = 1 : \ell$.

Proof. The existence of S is obvious as $\lambda_1, \ldots, \lambda_\ell$ are distinct eigenvalues of A. By Theorem 4.1.12(a), it follows that $\mathbf{M}(z) \simeq_{\lambda_j} CX_j(zI - A_j)^{-1}Y_j^*B$ is a minimal local realization of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. This proves (a).

Since $\mathbf{M}(z) \simeq_{\lambda_j} CX_j(zI - A_j)^{-1}Y_j^*B$ is a minimal local realization of $\mathbf{M}(z)$ at λ_j, it follows that $G_j(z) := CX_j(zI - A_j)^{-1}Y_j^*B$ is the principal part of the Laurent series of

$\mathbf{M}(z)$ at λ_j. Since $\mathrm{eig}(A_j) = \{\lambda_j\}$, the matrix $\lambda_j I - A$ is nilpotent. As the realization is minimal, the order of λ_j as a pole of $\mathbf{M}(z)$ is the same as the order of λ_j as a pole of $(zI - A_j)^{-1}$. Hence ν_j is the ascent of λ_j as an eigenvalue of A. Consequently,

$$\begin{aligned} G_j(z) &= CX_j(zI - A)^{-1}Y_j^* B = CX_j(z - \lambda_j + \lambda_j I - A)^{-1}Y_j^* B \\ &= \sum_{k=1}^{\nu_j - 1} \frac{CX_j(A - \lambda_j I)^k Y_j^* B}{(z - \lambda_j)^{k+1}} = P_j\left(\frac{1}{z - \lambda_j}\right) \text{ for } j = 1 : \ell. \end{aligned}$$

This shows that $P_j(z) = CX_j(z^{-1}I + \lambda_j I - A)^{-1}Y_j^* B = \sum_{k=1}^{\nu_j - 1} z^{k+1} CX_j(A - \lambda_j I)^k Y_j^* B$ is the principal pole polynomial of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. This proves (b) □

We have seen that a meromorphic matrix $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ admits a right coprime MFD of the form $\mathbf{M}(z) = N(z)D(z)^{-1}$, where $(N, D) \in \mathbb{H}(\Omega)^{m\times n} \times \mathbb{H}(\Omega)^{n\times n}$ and $D(z)$ is regular. If $\#(\wp_\Omega(\mathbf{M})) < \infty$ then we show that $D(z)$ can be chosen as a regular matrix polynomial. This fact is not evident from the Smith-McMillan form of $\mathbf{M}(z)$.

Theorem 4.1.16. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. Suppose that $\#(\wp_\Omega(\mathbf{M})) < \infty$. Then there exists $(N, D(z)) \in \mathbb{H}(\Omega)^{m\times n} \times \mathbb{C}[z]^{n\times n}$ with $D(z)$ being regular such that $N(z)$ and $D(z)$ are right coprime,*

$$\mathbf{M}(z) = N(z)D(z)^{-1} \text{ and } \wp_\Omega(\mathbf{M}) = \sigma_{\mathbb{C}}(D).$$

Proof. By Theorem 4.1.4, we have $\mathbf{M}(z) = \mathbf{H}(z) + G(z)$ and $\wp_\Omega(\mathbf{M}) = \wp_{\mathbb{C}}(G)$, where $\mathbf{H} \in \mathbb{H}(\Omega)^{m\times n}$ and $G(z) \in \mathbb{C}(z)^{m\times n}$ is a strictly proper rational matrix. Since $G(z)$ is a strictly proper rational matrix, we have a right coprime MFD $G(z) = \widehat{N}(z)D(z)^{-1}$ with $\wp_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(D)$, where $\widehat{N}(z) \in \mathbb{C}[z]^{m\times n}$ and $D(z) \in \mathbb{C}[z]^{n\times n}$ is regular; see [15, 32].

Set $N(z) := \mathbf{H}(z)D(z) + \widehat{N}(z)$. Then $N \in \mathbb{H}(\Omega)^{m\times n}$ and $\mathbf{M}(z) = N(z)D(z)^{-1}$ is a right MFD. Now

$$\begin{bmatrix} N(z) \\ D(z) \end{bmatrix} = \begin{bmatrix} \mathbf{H}(z)D(z) + \widehat{N}(z) \\ D(z) \end{bmatrix} = \begin{bmatrix} I_m & \mathbf{H}(z) \\ 0 & I_n \end{bmatrix} \begin{bmatrix} \widehat{N}(z) \\ D(z) \end{bmatrix}$$

shows that $\mathrm{rank}\left(\begin{bmatrix} N(z) \\ D(z) \end{bmatrix}\right) = \mathrm{rank}\left(\begin{bmatrix} \widehat{N}(z) \\ D(z) \end{bmatrix}\right) = n$ for all $z \in \Omega$. Hence $N(z)$ and $D(z)$ are right coprime. This shows that $\mathbf{M}(z) = N(z)D(z)^{-1}$ is a right coprime MFD. Now we have $\wp_\Omega(\mathbf{M}) = \wp_{\mathbb{C}}(G) = \sigma_{\mathbb{C}}(D)$. □

The next result shows that an irreducible analytic system matrix in which the state matrix is a regular matrix polynomial is strict system equivalent (Fuhrmann system equivalent) to an irreducible system matrix in SSF.

Theorem 4.1.17. *Let* $(C, B, D) \in \mathbb{H}(\Omega)^{m\times r} \times \mathbb{H}(\Omega)^{r\times n} \times \mathbb{H}(\Omega)^{m\times n}$ *and* $A(z) \in \mathbb{C}[z]^{r\times r}$ *be regular. Consider the system matrix*

$$\mathbf{H}(z) := \left[\begin{array}{c|c} A(z) & B(z) \\ \hline -C(z) & D(z) \end{array}\right]$$

and assume that $\mathbf{H}(z)$ *is irreducible. Then there exists a system matrix in SSF*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} zI_\ell - \widehat{A} & \widehat{B} \\ \hline -\widehat{C} & \widehat{D}(z) \end{array}\right], \text{ where } (\widehat{C}, \widehat{A}, \widehat{B}, \widehat{D}) \in \mathbb{C}^{m\times \ell} \times \mathbb{C}^{\ell\times\ell} \times \mathbb{C}^{\ell\times n} \times \mathbb{H}(\Omega)^{m\times n},$$

such that $\mathbf{S}(z)$ *is irreducible and* $\mathbf{H}(z) \sim_{fse} \mathbf{S}(z)$.

Proof. Note that $\mathbf{M}(z) := D(z)+C(z)A(z)^{-1}B(z)$ is the transfer of $\mathbf{H}(z)$ and $\wp_\Omega(\mathbf{M}) = \sigma_{\mathbb{C}}(A)$. Since $A(z)$ is a regular matrix polynomial, we have $\#(\wp_\Omega(\mathbf{M})) < \infty$. Hence by Theorem 4.1.14, an irreducible system matrix in SSF such as $\mathbf{S}(z)$ exists having the transfer function $\mathbf{M}(z)$. Since $\mathbf{H}(z)$ and $\mathbf{S}(z)$ are both irreducible and have the same transfer function, by Theorem 2.4.16, we have $\mathbf{H}(z) \sim_{fse} \mathbf{S}(z)$. □

We now summarize the results for $\mathbf{M}(z)$ when $\#(\wp_\Omega(\mathbf{M})) < \infty$.

Theorem 4.1.18. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$. *Suppose that* $\#(\wp_\Omega(\mathbf{M})) < \infty$. *Then we have the following.*

(a) There exists $(C, A, B) \in \mathbb{C}^{m\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$, *for some* $r \in \mathbb{N}$, *and* $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{m\times n}$ *such that* $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ *is a minimal local realization of* $\mathbf{M}(z)$ *on* Ω *and* $\wp_\Omega(\mathbf{M}) = \mathrm{eig}(A)$. *The system matrix in SSF*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right]$$

is irreducible and $\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{S})$.

(b) There exists $(N, D(z)) \in \mathbb{H}(\Omega)^{m\times n} \times \mathbb{C}[z]^{n\times n}$ *with* $D(z)$ *being a regular matrix polynomial such that* $\wp_\Omega(\mathbf{M}) = \sigma_{\mathbb{C}}(D)$ *and* $\mathbf{M}(z) = N(z)D(z)^{-1}$ *is a right coprime MFD. The RMF-system matrix*

$$\mathbf{S}_R(z) := \begin{bmatrix} D(z) & I_n \\ -N(z) & 0_{m\times n} \end{bmatrix}$$

is irreducible and $\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{S}_R)$.

(c) *We have* $\mathbf{S}(z) \sim_{fse} \mathbf{S}_R(z)$. *Let* $p \geq \max(r, n)$. *Then the following hold:*

(i) $I_{p-r} \oplus \mathbf{S}(z) \sim_{sse} I_{p-n} \oplus \mathbf{S}_R(z) \sim_\Omega I_p \oplus N(z)$.

(ii) $I_{p-r} \oplus (A - zI_r) \sim_\Omega I_{p-n} \oplus D(z)$.

(d) *For the Smith forms, we have*

$$I_{p-r} \oplus S_A(z) = I_{p-n} \oplus S_D(z) \text{ and } I_{p-r} \oplus S_{\mathbf{S}}(z) = I_{p-n} \oplus S_{\mathbf{S}_R}(z) = I_p \oplus S_N(z),$$

where $S_A(z)$ *denotes the Smith form of the matrix* $A - zI_r$ *and* $S_X(z)$ *denotes the Smith form of a holomorphic matrix* $X(z)$. *In particular, we have*

$$\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{S}) = \sigma_\Omega(\mathbf{S}_R) = \sigma_\Omega(N) \text{ and } \wp_\Omega(\mathbf{M}) = \mathrm{eig}(A) = \sigma_\Omega(D).$$

Further, $\mathrm{ind}_e(\lambda, \mathbf{M}) = \mathrm{ind}_e(\lambda, \mathbf{S}) = \mathrm{ind}_e(\lambda, \mathbf{S}_R) = \mathrm{ind}_e(\lambda, N)$ *for* $\lambda \in \sigma_\Omega(\mathbf{M})$ *and* $\mathrm{ind}_p(\lambda, \mathbf{M}) = \mathrm{ind}_e(\lambda, A) = \mathrm{ind}_e(\lambda, D)$ *for* $\lambda \in \wp_\Omega(\mathbf{M})$.

Proof. The results in (a) follow from Theorem 4.1.14 and the results in (b) follow from Theorem 4.1.16 and Theorem 2.4.19. The results in (c) and (d) follow from Theorem 2.4.19. □

4.1.1 Computation of local realization by contour integration

Let $A \in \mathbb{C}^{n\times n}$ and $f : \mathbb{C} \longrightarrow \mathbb{C}$ be analytic. Let Γ be a positively oriented rectifiable simple closed curve in $\mathbb{C}$ such that $\mathrm{eig}(A) \subset \mathrm{Int}(\Gamma)$, where $\mathrm{Int}(\Gamma)$ denotes the region enclosed by the curve Γ. Then we have

$$f(A) = \frac{1}{2\pi i}\int_\Gamma f(z)(zI_n - A)^{-1}dz. \tag{4.2}$$

Theorem 4.1.19. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{m\times n}$ *and* $\mathbf{M}(z) \simeq_\Omega C(zI_r - A)^{-1}B$ *be a minimal local realization. Let* $\Gamma \subset \Omega$ *be a positively oriented rectifiable simple closed curve in* Ω *such that* $\wp_\Omega(\mathbf{M}) \subset \mathrm{Int}(\Gamma) \subset \Omega$. *Then for* $f \in \mathbb{H}(\Omega)$ *we have*

$$\frac{1}{2\pi i}\int_\Gamma f(z)\mathbf{M}(z)dz = \frac{1}{2\pi i}\int_\Gamma f(z)C(zI_r - A)^{-1}Bdz = Cf(A)B.$$

In particular, we have $S_\ell := \dfrac{1}{2\pi i}\displaystyle\int_\Gamma z^\ell \mathbf{M}(z)dz = CA^\ell B$ *for* $\ell \in \mathbb{Z}_+$.

Proof. There exists $\mathbf{H} \in \mathbb{H}(\Omega)^{m\times n}$ such that $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$. Since $f(z)\mathbf{H}(z)$ is holomorphic in Ω, by Cauchy's theorem $\int_\Gamma f(z)\mathbf{H}(z)dz = 0$. Hence by (4.2),

$$\frac{1}{2\pi i}\int_\Gamma f(z)\mathbf{M}(z)dz = \frac{1}{2\pi i}\int_\Gamma f(z)C(zI_r - A)^{-1}Bdz = Cf(A)B.$$

For $f_\ell(z) := z^\ell$, we have $S_\ell = CA^\ell B$ for $\ell \in \mathbb{Z}_+$. □

The matrices S_ℓ for $\ell \in \mathbb{Z}_+$ are called Markov parameters of $\mathbf{M}(z)$ on Ω or the Markov parameters of $G_{sp}(z) := C(zI_r - A)^{-1}B$. Computation of (C, A, B) from the Markov parameters is a well studied topic in realization theory. There are several algorithms for computation of (C, A, B) from the Markov parameters; see [15]. We briefly describe an algorithm based on orthogonal matrix factorization. Consider the block Hankel matrix $\mathbb{H}_\ell$ and the shifted block Hankel matrix $\widehat{\mathbb{H}}_\ell$ given by

$$\mathbb{H}_\ell := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \\ \vdots & \ddots & \vdots \\ S_{\ell-1} & \cdots & S_{2\ell-2} \end{bmatrix} \text{ and } \widehat{\mathbb{H}}_\ell := \begin{bmatrix} S_1 & \cdots & S_\ell \\ \vdots & \ddots & \vdots \\ S_\ell & \cdots & S_{2\ell-1} \end{bmatrix} \text{ for } \ell \in \mathbb{Z}_+.$$

Then it follows that

$$\mathbb{H}_\ell = \begin{bmatrix} CB & CAB & \cdots & CA^{\ell-1}B \\ CAB & CA^2B & \cdots & CA^{\ell+1}B \\ \vdots & \vdots & \vdots & \vdots \\ CA^{\ell-1}B & CA^\ell B & \cdots & CA^{2\ell-2}B \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix} \begin{bmatrix} I_r & AB & \dots & A^{\ell-1}B \end{bmatrix},$$

$$\widehat{\mathbb{H}}_\ell = \begin{bmatrix} CAB & CA^2B & \cdots & CA^\ell B \\ CA^2B & CA^3B & \cdots & CA^{\ell+2}B \\ \vdots & \vdots & \vdots & \vdots \\ CA^\ell B & CA^{\ell+1}B & \cdots & CA^{2\ell-1}B \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix} A \begin{bmatrix} I_r & AB & \dots & A^{\ell-1}B \end{bmatrix}.$$

This shows that $\operatorname{rank}(\mathbb{H}_\ell) \leq r$ for $\ell \in \mathbb{N}$ and there exists $\nu \in \mathbb{N}$ such that $\operatorname{rank}(\mathbb{H}_\ell) = r$ for all $\ell \geq \nu$. The index ν is in fact the minimality index of the observable pair (C, A) and the controllable pair (A, B); see [15]. In particular, we have $\operatorname{rank}(\mathbb{H}_\ell) = r$ for $\ell \geq r$. It also follows that $\operatorname{rank}(\widehat{\mathbb{H}}_\ell) \leq r$ for $\ell \in \mathbb{N}$

Consider the controllability matrix $\mathbf{C}_\ell$ and the observablity matrix $\mathbf{O}_\ell$ given by

$$\mathbf{C}_\ell := \begin{bmatrix} B & AB & \cdots & A^{\ell-1}B \end{bmatrix} \text{ and } \mathbf{O}_\ell := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix}. \tag{4.3}$$

Then we have $\mathbb{H}_\ell = \mathbf{O}_\ell \mathbf{C}_\ell$ and $\widehat{\mathbb{H}}_\ell = \mathbf{O}_\ell A \mathbf{C}_\ell$ for $\ell \in \mathbb{N}$. Thus,

$$\operatorname{rank}(\mathbb{H}_\ell) = \operatorname{rank}(\mathbf{O}_\ell) = \operatorname{rank}(\mathbf{C}_\ell) = r \quad \text{for} \quad \ell \geq r.$$

Next, observe that for $\ell > r$ we have

$$\mathbf{O}_{\ell-1}A = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-2} \end{bmatrix} A = \begin{bmatrix} CA \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix} = \mathbf{O}_\ell(2:\ell,:) \Longrightarrow A = \mathbf{O}_{\ell-1}^\dagger \mathbf{O}_\ell(2:\ell,:)$$

and

$$A\mathbf{C}_{\ell-1} = A\begin{bmatrix} B & AB & \cdots & A^{\ell-2}B \end{bmatrix} = \begin{bmatrix} AB & A^2B & \cdots & A^{\ell-1}B \end{bmatrix} = \mathbf{C}_\ell(:,2:\ell)$$

which shows that $A = \mathbf{C}_\ell(:,2:\ell)\mathbf{C}_{\ell-1}^\dagger$. Here $X^\dagger$ is the Moore-Penrose pseudo-inverse of X, $\mathbf{O}_\ell(2:\ell,:)$ is the block sub-matrix containing block rows 2 to ℓ of $\mathbf{O}_\ell$ and $\mathbf{C}_\ell(:,2:\ell)$ is the block sub-matrix containing block columns $2:\ell$ of $\mathbf{C}_\ell$. The observations above yield a method for computing a minimal realization (C, A, B) from the observability matrix $\mathbf{O}_\ell$ and the controllability matrix $\mathbf{C}_\ell$.

Let $\mathbb{H}_\ell = UV^*$ be an orthogonal factorization, where $U \in \mathbb{C}^{m\ell\times r}$ and $V \in \mathbb{C}^{n\ell\times r}$ have orthogonal columns and full column rank. Since $UV^* = \mathbb{H}_\ell = \mathbf{O}_\ell\mathbf{C}_\ell$, we can choose $\mathbf{O}_\ell := UT$ and $\mathbf{C}_\ell := T^{-1}V^*$ for any nonsingular matrix $T \in \mathbb{C}^{r\times r}$. We choose $T = I_r$. Then the relevant block rows of $\mathbf{O}_\ell$ are given by $\mathbf{O}_{\ell-1} = U(1:(\ell-1)m, :)$ and $\mathbf{O}_\ell(2:\ell, :) = U(m+1:m\ell, :)$. This yields a minimal realization (C, A, B) :

$$C := U(1:m, :),\ A := U(1:(\ell-1)m, :)^\dagger U(m+1:m\ell, :),\ B := V^*(1:n, :).$$

Note that C is the first m rows of $\mathbf{O}_\ell = U$ and B is the first n columns of $\mathbf{C}_\ell = V^*$. Hence $C = (e_1^\top \otimes I_m)U$ and $B = V^*(e_1 \otimes I_n)$, where e_1 is the first column of I_ℓ. The matrix A can also be determined from $\widehat{\mathbb{H}}_\ell$ as follows:

$$\widehat{\mathbb{H}}_\ell = \mathbf{O}_\ell A\mathbf{C}_\ell \Longrightarrow A = \mathbf{O}_\ell^\dagger\, \widehat{\mathbb{H}}_\ell\, \mathbf{C}_\ell^\dagger = U^\dagger\, \widehat{\mathbb{H}}_\ell\, (V^*)^\dagger.$$

The orthogonal factors U and V can be computed from a compact SVD of $\mathbb{H}_\ell$. Indeed, consider the compact SVD $\mathbb{H}_\ell = U\Sigma_r V^*$, where $\Sigma_r := \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ and $\sigma_1, \ldots, \sigma_r$ are the nonzero singular values of $\mathbb{H}_\ell$. Then considering

$$\mathbf{O}_\ell := U\Sigma^{1/2} \text{ and } \mathbf{C}_\ell := \Sigma^{1/2}V^*$$

we have

$$C := (e_1^\top \otimes I_m)\mathbf{O}_\ell,\ A := \Sigma^{-1/2}U^*\, \widehat{\mathbb{H}}_\ell\, V\Sigma^{-1/2},\ \ B := \mathbf{C}_\ell(e_1 \otimes I_n).$$

This yields the following SVD based algorithm; see [15, Chapter 9, Algorithm 9.3.2].

Algorithm-A(minimal realization).

Input: $\mathbb{H}_\ell$ and $\widehat{\mathbb{H}}_\ell$.
Output: minimal realization (C, A, B)

1. Compute compact SVD $\mathbb{H}_\ell = U\Sigma V^*$.
2. Compute $C := (e_1^\top \otimes I_m)U\Sigma^{1/2} = (U\Sigma^{1/2})(1 : m, :)$.
3. Compute $B := \Sigma^{1/2}V^*(e_1 \otimes I_n) = (\Sigma^{1/2}V^*)(:, 1 : n)$.
4. Compute $A := \Sigma^{-1/2}U^*\widehat{\mathbb{H}}_\ell V\Sigma^{-1/2}$.

Alternatively, we can compute (C, A, B) using only $\mathbb{H}_\ell$. We have the following algorithm in MATLAB notation.

Algorithm-B(minimal realization).

Input: $\mathbb{H}_\ell$.
Output: minimal realization (C, A, B)

1. `[U, S, V] = svd(`$\mathbb{H}_\ell$`, 0); % compact SVD`
2. `U = U*sqrt(S); V = V*sqrt(S); % orthogonal factors`
3. `C = U(1:m, :); B = V'(:, 1:n);`
4. `A = U(1:(ℓ-1)m, :)\U(m+1:mℓ, :); % least-squares solution or`
   `A = pinv(U(1:(ℓ-1)m, :))*U(m+1:mℓ, :); % apply pseudo-inverse`

We mention that Algorithm-B is known to perform poorly for rational matrices.

4.1.2 Solution of nonlinear eigenvalue problems

Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular, that is, $\mathbf{T}(z)$ is invertible for some $z \in \Omega$. Consider the nonlinear eigenvalue problem (NEP): Find $\lambda \in \Omega$ and a nonzero $v \in \mathbb{C}^n$ such that

$$\mathbf{T}(\lambda)v = 0. \tag{4.4}$$

The scalar λ is called an eigenvalue of $\mathbf{T}(z)$ and v is called an eigenvector of $\mathbf{T}$ corresponding to λ. NEPs arise in many applications. We have seen that the NEP

$$\mathbf{T}(\lambda)v := (\lambda I - A_0 - \sum_{i=1}^{m} A_i e^{-\lambda\tau_i})v = 0$$

arises in the study of the delay differential equation

$$\frac{d}{dt}x(t) = A_0 x(t) + \sum_{i=1}^{m} A_i x(t-\tau_i),$$

where $x(t) \in \mathbb{R}^n$ is the state variable at time t, A_i's are $n \times n$ matrices, and $0 < \tau_1 < \tau_2 < \cdots < \tau_m$ represent the time-delays [45].

It is a challenging task to solve an NEP. However, eigenvalues can be treated as poles of appropriate meromorphic matrices and minimal local realization of meromorphic matrices can be utilized for solving an NEP. Observe that $\mathbf{M}(s) := \mathbf{T}(z)^{-1}$ is meromorphic in Ω and

$$\lambda \in \sigma_\Omega(\mathbf{T}) \Longleftrightarrow \lambda \in \wp_\Omega(\mathbf{M}).$$

Hence finding eigenvalues of $\mathbf{T}(z)$ is equivalent to finding poles of $\mathbf{M}(z)$. Thus, any pole finding method for $\mathbf{M}(z)$ can be employed to find the eigenvalues of $\mathbf{T}(z)$. As a consequence of Theorem 4.1.14, we have the following result.

Theorem 4.1.20. *Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular and $\mathcal{O} \subset \Omega$ be open. Suppose that $\#(\sigma_\mathcal{O}(\mathbf{T})) < \infty$. Then there exists $(C, A, B) \in \mathbb{C}^{m\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$, for some $r \in \mathbb{N}$, and $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{n\times n}$ such that $\mathbf{T}(z)^{-1} = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ is a minimal local realization on $\mathcal{O}$ and $\sigma_\mathcal{O}(\mathbf{T}) = \mathrm{eig}(A) = \sigma_\Omega(\mathbf{T}) \cap \mathcal{O}$.*

Let $\mathcal{O} \subset \Omega$ be an open and $\mathbf{T}(z)^{-1} \simeq_\mathcal{O} C(zI_r - A)B$ be a minimal local realization. Then it follows that $\sigma(A) \subset \sigma_\Omega(\mathbf{T})$. Further, it can be shown that Cv is an eigenvector of $\mathbf{T}(z)$ whenever v is an eigenvector of A. Hence a minimal local realization can be utilized for solving a nonlinear eigenvalue problem.

Theorem 4.1.21. *Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular and $\mathcal{O} \subset \Omega$ be open. Suppose that $\mathbf{T}(z)^{-1} \simeq_\mathcal{O} C(zI_r - A)B$ is a minimal local realization. Then $\sigma_\Omega(\mathbf{T}) \cap \mathcal{O} = \mathrm{eig}(A)$ and*

$$Av = \lambda v \text{ and } v \neq 0 \Longrightarrow \mathbf{T}(\lambda)Cv = 0 \text{ and } Cv \neq 0.$$

Further, the partial multiplicities of λ as an eigenvalue of A are the same as the partial multiplicities of λ as an eigenvalue of $\mathbf{T}(z)$.

Proof. Obviously $\text{eig}(A) \subset \sigma_\Omega(\mathbf{T})$ and $\sigma_\Omega(\mathbf{T}) \cap \mathcal{O} = \text{eig}(A)$. Let Γ be a simple closed curve in $\mathcal{O}$ such that $\sigma_\Omega(\mathbf{T}) \cap \text{Int}(\Gamma) = \text{eig}(A)$. Now $Av = \lambda v$ and $v \neq 0 \Longrightarrow v$ is an eigenvector of A corresponding to the eigenvalue λ. Since (C, A) is observable, we have

$$\text{rank}\left(\begin{bmatrix} \lambda I_r - A \\ C \end{bmatrix}\right) = r \Longrightarrow Cv \neq 0.$$

Since $(zI_r - A)^{-1}v = v/(z-\lambda)$, by Cauchy's integral formula, we have

$$\mathbf{T}(\lambda)Cv = \frac{1}{2\pi i}\int_\Gamma \mathbf{T}(z)C(zI_r - A)^{-1}v dz = 0$$

whenever $S := \int_\Gamma \mathbf{T}(z)C(zI_r - A)^{-1}dz = 0$. We show that $S = 0$ which follows from the controllability of (A, B). Indeed, since (A, B) is controllable, the controllability matrix $\mathbf{C}_r := \begin{bmatrix} B & AB & \cdots & A^{r-1}B \end{bmatrix}$ has full row rank. Hence $S = 0 \Longleftrightarrow S\mathbf{C}_r = 0$.

We show that $S\mathbf{C}_r = \int_\Gamma \mathbf{T}(z)C(zI_r - A)^{-1}\begin{bmatrix} B & AB & \cdots & A^{r-1}B \end{bmatrix} dz = 0$. First, note that $(zI_r - A)^{-1}A^j = (zI_r - A)^{-1}(A - zI_r + zI_r)^j \simeq_{\mathcal{O}} z^j(zI_r - A)^{-1}$ for $j = 0 : r-1$. This shows that $\mathbf{T}(z)C(zI_r - A)^{-1}A^jB \simeq_{\mathcal{O}} z^j\mathbf{T}(z)C(zI_r - A)^{-1}B$ for $j = 0 : r-1$. Next, note that $\mathbf{T}(z)^{-1} - C(zI_r - A)^{-1}B$ is analytic in $\mathcal{O}$ which implies that $\mathbf{T}(z)C(zI_r - A)^{-1}B$ is analytic in $\mathcal{O}$ which in turn implies that $z^j\mathbf{T}(z)C(zI_r - A)^{-1}B$ is analytic in $\mathcal{O}$ which implies that $\int_\Gamma z^j\mathbf{T}(z)C(zI_r - A)^{-1}Bdz = 0$ for $j = 0 : r-1$. Hence we have $\int_\Gamma \mathbf{T}(z)C(zI_r - A)^{-1}A^jBdz = \int_\Gamma z^j\mathbf{T}(z)C(zI_r - A)^{-1}Bdz = 0$ for $j = 0 : r-1$. This shows that $S\mathbf{C}_r = 0$.

Finally, by Theorem 4.1.18, the partial multiplicities of λ as an eigenvalue of $\mathbf{T}(z)$ are the same as the partial multiplicities of λ as a pole of $\mathbf{T}(z)^{-1}$. On the other hand, the partial multiplicities of λ as a pole of $\mathbf{T}(z)^{-1}$ are the same as the partial multiplicities of λ as an eigenvalue of A. This completes the proof. □

Remark 4.1.22. *Let $\lambda \in \sigma_\Omega(\mathbf{T})$. Then a nonzero $u \in \mathbb{C}^n$ is called a left eigenvector of $\mathbf{T}(z)$ corresponding to λ if $u^\top\mathbf{T}(\lambda) = 0$, that is, $\mathbf{T}(\lambda)^\top u = 0$. Now if*

$$\mathbf{T}(z)^{-1} \simeq_{\mathcal{O}} C(zI_r - A)B$$

is a minimal local realization then $\mathbf{T}(z)^{-\top} \simeq_{\mathcal{O}} B^\top(zI_r - A^\top)C^\top$ is a minimal local realization. Hence by Theorem 5.2.10 we have the following result:

$$A^\top u = \lambda u \text{ and } u \neq 0 \Longrightarrow \mathbf{T}(\lambda)^\top B^\top u = 0 \text{ and } B^\top u \neq 0.$$

Thus, if u is a left eigenvector of A then $B^\top u$ is a left eigenvector of $\mathbf{T}(z)$.

Keldysh [34] obtained a representation of $\mathbf{T}(z)^{-1}$, which is similar to a minimal local realization in Theorem 4.1.20, involving generalized eigenvectors and a ***Jordan matrix***, that is, a matrix in Jordan canonical form; see [44, Theorem 1.6.5]. A proof of Keldysh theorem can be found in [44] which uses canonical system of root functions. See also [29, 35, 11].

Theorem 4.1.23 (Keldysh [44]). *Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular and $\mathcal{O} \subset \Omega$ be open. Suppose that $\sigma_{\mathcal{O}}(\mathbf{T}) = \{\lambda_1, \ldots, \lambda_\ell\}$ and the total algebraic multiplicity is m. Then there exist $V, W \in \mathbb{C}^{n\times m}$, a Jordan matrix $J \in \mathbb{C}^{m\times m}$ and $\mathbf{H} \in \mathbb{H}(\mathcal{O})^{n\times n}$ such that*

$$\mathbf{T}^{-1}(z) = \mathbf{H}(z) + V(zI_m - J)^{-1}W^* \text{ and } \sigma_{\mathcal{O}}(\mathbf{T}) = \operatorname{eig}(J).$$

The partial multiplicities of λ as an eigenvalue of J are the same as the partial multiplicities of λ as an eigenvalue of $\mathbf{T}(z)$. The columns of V and W are right and left generalized eigenvectors of $\mathbf{T}(z)$ and satisfy certain normalization constraints.

Let $\lambda \in \sigma_\Omega(\mathbf{T})$ and m be the algebraic multiplicity of λ. Let $J \in \mathbb{C}^{m\times m}$ be a uni-spectral Jordan matrix with $\sigma(J) = \{\lambda\}$ and $X \in \mathbb{C}^{n\times m}$. Then (X, J) is said to be a ***Jordan pair*** of $\mathbf{T}(z)$ at λ if (X, J) is observable and $\mathbf{T}(z)X(zI - J)^{-1}$ is analytic at λ. Gohberg et al. [24] developed a local spectral theory for $\mathbf{T}(z)$ using canonical system of root functions as the main tools to study Jordan pairs of $\mathbf{T}(z)$. It is shown [24] that a Jordan pair of $\mathbf{T}(z)$ at λ exists and can be constructed from a canonical system of root functions at λ. A Jordan pair at λ provides a local representation of $\mathbf{T}(z)^{-1}$ which is similar in spirit to the representation in Theorem 4.1.20.

Theorem 4.1.24 (Gohberg et al.,[24])). *Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular. Let $\lambda \in \sigma_\Omega(\mathbf{T})$ and m be the algebraic multiplicity of λ. Let $(X, J) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{m\times m}$ be such that (X, J) is observable and J is a Jordan matrix with $\operatorname{eig}(J) = \{\lambda\}$. Then (X, J) is a Jordan pair of $\mathbf{T}(z)$ at λ if and only if there exists $Y \in \mathbb{C}^{m\times n}$ such that (J, Y) is controllable and*

$$\mathbf{T}(z)^{-1} \simeq_\lambda X(zI_m - J)^{-1}Y$$

Let $\mathcal{O} \subset \Omega$ be open. Suppose that $\sigma_{\mathcal{O}}(\mathbf{T}) = \{\lambda_1, \ldots, \lambda_\ell\}$ and the total algebraic multiplicity is m. Then it follows from Theorem 4.1.24 that there exists $(X, J, Y) \in \mathbb{C}^{n\times m} \times \mathbb{C}^{m\times m} \times \mathbb{C}^{m\times n}$ such that J is a Jordan matrix with $\operatorname{eig}(J) = \{\lambda_1, \ldots, \lambda_\ell\}$, (X, J) is observable, (Y, B) is controllable and

$$\mathbf{T}(z)^{-1} \simeq_{\mathcal{O}} X(zI_m - J)^{-1}Y. \tag{4.5}$$

We mention that Theorem 4.1.23 and Theorem 4.1.24 as well as the representation in (4.5) follow as special cases from Theorem 4.1.20 by reducing A to Jordan canonical form by similarity transformation and using the result in Theorem 4.1.12. Note that Theorem 4.1.20 is a special case of Theorem 4.1.14. The local realization of $\mathbf{T}(z)^{-1}$ can be computed using either Algorithm-A or Algorithm-B with $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$ which results in a contour integral based method for solving the NEP $\mathbf{T}(\lambda)v = 0$.

Contour integral based methods for solving $\mathbf{T}(\lambda)v = 0$ have received a lot of attention in recent time; see [29, 11, 5, 60, 21, 12] and the references therein. All these methods are directly or indirectly related to a minimal local realization of $\mathbf{T}(z)^{-1}$. All these methods project $\mathbf{T}(z)^{-1}$ on to a lower dimensional space so as to reduce the cost of computation.

The contour integral based methods can be formulated in a unified framework as follows. Let $\mathbf{L} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{\ell\times p}$ be a linear transformation, where $\max(\ell, p) \ll n$. Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular. Now consider $\mathbf{M}(z) := \mathbf{L}(\mathbf{T}(z)^{-1})$. Then $\mathbf{M}(z)$ is meromorphic and $\wp_\Omega(\mathbf{M}) \subset \sigma_\Omega(\mathbf{T})$. Therefore, in view of Theorem 4.1.14 and Theorem 4.1.19, a subset of $\sigma_\Omega(\mathbf{T})$ can be computed by considering a minimal local realization of $\mathbf{M}(z)$ and using Algorithm-A or Algorithm-B. Thus as a consequence of Theorem 4.1.14 and Theorem 4.1.19, we have the following result.

Theorem 4.1.25. *Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular and $\mathbf{L} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{\ell\times p}$ be a linear transformation, where $\max(\ell, p) \ll n$. Define $\mathbf{M}(z) := \mathbf{L}(\mathbf{T}(z)^{-1})$. Then $\mathbf{M} \in \mathbb{M}(\Omega)^{\ell\times p}$ and $\wp_\Omega(\mathbf{M}) \subset \sigma_\Omega(\mathbf{T})$. Let $\mathcal{O} \subset \Omega$ be open. Suppose that $\#(\wp_{\mathcal{O}}(\mathbf{M})) < \infty$. Then there exist $(C, A, B) \in \mathbb{C}^{\ell\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times p}$, for some $r \in \mathbb{N}$, such that*

$$\mathbf{M}(z) \simeq_{\mathcal{O}} C(zI_r - A)^{-1}B$$

is a minimal local realization of $\mathbf{M}(z)$ on $\mathcal{O}$. Let Γ be a simple closed curve in $\mathcal{O}$ such that $\wp_{\mathcal{O}}(\mathbf{M}) \subset \mathrm{Int}(\Gamma) \subset \mathcal{O}$. Consider the matrix moments

$$S_k := \frac{1}{2\pi i}\int_\Gamma z^k \mathbf{M}(z)dz = \frac{1}{2\pi i}\int_\Gamma z^k \mathbf{L}(\mathbf{T}(z)^{-1})dz, \quad \textit{for} \quad k \in \mathbb{Z}_+,$$

and construct the block Hankel matrix $\mathbb{H}_m$ and shifted block Hankel matrix $\widehat{\mathbb{H}}_m$ for large enough m. Then (C, A, B) can be constructed from $\mathbb{H}_m$ and $\widehat{\mathbb{H}}_m$ by Algorithm-A and Algorithm-B.

The contour integral based method proposed by Beyn [11] considers an arbitrarily chosen matrix $V \in \mathbb{C}^{n\times p}$ with full column rank and the matrix moment

$$S_k := \frac{1}{2\pi i}\int_\Gamma z^k \mathbf{T}(z)^{-1} V dz, \quad \text{for} \quad k \in \mathbb{Z}_+.$$

These moments correspond to the linear transformation $\mathbf{L} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times p}$ given by $\mathbf{L}(X) := XV$, where $V \in \mathbb{C}^{n\times p}$ is a full column rank matrix. On the other hand, for arbitrary $u, v \in \mathbb{C}^n$, Asakura et al. [5] consider the scalar moments

$$s_k := \frac{1}{2\pi i}\int_\Gamma z^k u^\top \mathbf{T}(z)^{-1} v dz, \quad \text{for} \quad k \in \mathbb{Z}_+,$$

which correspond to the linear transformation $\mathbf{L} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}$ given by $\mathbf{L}(X) := u^\top X v$, where $u, v \in \mathbb{C}^n$ are fixed vectors. Other choices of $\mathbf{L}$ such as $\mathbf{L}(X) := UXV$, where $U \in \mathbb{C}^{\ell\times n}$ has full row rank and $V \in \mathbb{C}^{n\times p}$ has full column rank, have also been considered in the literature.

Remark 4.1.26. *Let $\mathbf{T}(z)^{-1} \simeq_{\mathcal{O}} C(zI_r - A)^{-1}B$ be a minimal local realization. Let $V \in \mathbb{C}^{n\times p}$ be a full column rank matrix such that (A, BV) is controllable. Then*

$$\mathbf{T}(z)^{-1}V \simeq_{\mathcal{O}} C(zI_r - A)^{-1}BV$$

is a minimal local realization which corresponds to the transformation $\mathbf{L}(X) := XV$. In this case, Theorem 5.2.10 still holds and we have the following result:

$$Av = \lambda v \text{ and } v \neq 0 \Longrightarrow \mathbf{T}(\lambda)Cv = 0 \text{ and } Cv \neq 0.$$

4.2 Linearization of a meromorphic matrix

Linearization is a classical technique widely used for solving polynomial eigenvalue problems. Let $P(z) \in \mathbb{C}[z]^{n\times n}$ be a matrix polynomial of degree m. Then the polynomial eigenvalue problem (PEP) solves

$$P(\lambda)v = 0$$

for $\lambda \in \mathbb{C}$ and a nonzero $v \in \mathbb{C}^n$. The scalar λ is called an eigenvalue of $P(z)$ and v is called an eigenvector of $P(z)$ corresponding to λ. The PEP can be transformed to an equivalent generalized eigenvalue problem (GEP)

$$(\lambda X + Y)\mathbf{v} = 0, \text{ where } (\lambda, \mathbf{v}) \in \mathbb{C} \times \mathbb{C}^{mn},$$

by means of linearization of $P(z)$ via unimodular equivalence; see [25, 41] and the references therein. A matrix polynomial $U(z) \in \mathbb{C}[z]^{n\times n}$ is unimodular if $\det(U(z))$ is a nonzero constant independent of z.

Definition 4.2.1 ([25, 41])**.** *Let $P(z) \in \mathbb{C}[z]^{n\times n}$ be a matrix polynomial of degree m. A matrix pencil $L(z) := zX + Y$ with $X, Y \in \mathbb{C}^{mn\times mn}$ is called a linearization of $P(z)$ if there exist $mn \times mn$ unimodular matrix polynomials $U(z)$ and $V(z)$ such that*

$$U(z)L(z)V(z) = \left[\begin{array}{c|c} I_{(m-1)n} & 0 \\ \hline 0 & P(z) \end{array}\right].$$

Eigenvalue problems involving rational matrices also arise in many applications such as in acoustic emissions of high speed trains, calculation of quantum dots, free vibration of plates with elastically attached masses, vibration of fluid solid structure and in control theory; see [40, 10] and the references therein. Let $G(z) \in \mathbb{C}(z)^{n\times n}$ be a rational matrix. Then the rational eigenvalue problem (REP) solves

$$G(\lambda)v = 0$$

for $\lambda \in \mathbb{C}$ and a nonzero $v \in \mathbb{C}^n$. The scalar λ is called an eigenvalue of $G(z)$ and v is called an eigenvector of $G(z)$ corresponding to λ. For example, the REP

$$G(\lambda)v := -Av + \lambda Bv + \lambda^2 \sum_{j=1}^{k} \frac{1}{w_j - \lambda} C_j v = 0$$

arises when a generalized linear eigenvalue problem condensed exactly; see [43] for details.

With a view to transforming an REP to an equivalent GEP, a concept of linearization of rational matrices has been developed recently; see [2, 3, 14, 53, 4] and the references therein. Linearization of a rational matrix has been defined in two different ways, namely, via MFD and state-space realization; see [2].

Let $G_1(z), G_2(z) \in \mathbb{C}(z)^{m\times n}$. Then $G_1(z)$ is said to be unimodularly equivalent to $G_2(z)$ and written as $G_1(z) \sim G_2(z)$ if there exist unimodular matrix polynomials $U(z) \in \mathbb{C}[z]^{m\times m}$ and $V(z) \in \mathbb{C}[z]^{n\times n}$ such that $G_2(z) = U(z)G_1(z)V(z)$. If $G(z) \in \mathbb{C}(z)^{m\times n}$ is a rational matrix then there exists a matrix polynomial $P(z) \in \mathbb{C}[z]^{m\times n}$ and a strictly proper rational matrix $G_{sp}(z) \in \mathbb{C}(z)^{m\times n}$ such that $G(z) = P(z) + G_{sp}(z)$. Define $\deg(G(z)) := \deg(P(z))$, the degree of the polynomial part of $G(z)$.

Definition 4.2.2. *[2] Let $G(z) \in \mathbb{C}[z]^{n\times n}$ be a rational matrix and $G(z) = N(z)D(z)^{-1}$ be a right coprime MFD. Set $r := \deg\det(D(z))$ and $m := \deg(G(z))$. Then an $(nm + r) \times (nm + r)$ matrix pencil of the form*

$$\mathbb{L}(z) := \left[\begin{array}{c|c} X + zY & C \\ \hline B & A - zI_r \end{array}\right]$$

is called a linearization of $G(z)$ if

$$\left[\begin{array}{c|c} I_{s-n} & 0 \\ \hline 0 & N(z) \end{array}\right] \sim \left[\begin{array}{c|c} I_{s-mn-r} & 0 \\ \hline 0 & \mathbb{L}(z) \end{array}\right] \text{ and } \left[\begin{array}{c|c} I_{p-n} & 0 \\ \hline 0 & D(z) \end{array}\right] \sim \left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & A - zI_r \end{array}\right],$$

where $p := \max(r, n)$ and $s := \max(mn + r, mn + n)$.

If $G(z)$ is a matrix polynomial then $D(z) = I$ and hence $r = 0$. In such a case, $\mathbb{L}(z) = X + zY$ reduces to a linearization of the matrix polynomial $G(z)$. Alternatively, a minimal realization of $G(z)$ can be used to define a linearization of the system matrix which is equivalent to a linearization of $G(z)$.

Definition 4.2.3. *[2] Let $G(z) = P(z) + C(zI_r - A)^{-1}B$ be a minimal realization of $G(z) \in \mathbb{C}(z)^{n\times n}$, where $(C, A, B) \in \mathbb{C}^{n\times r} \times \mathbb{C}^{r\times r} \times \mathbb{C}^{r\times n}$ and $P(z) \in \mathbb{C}[z]^{n\times n}$. Consider the system matrix in SSF*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} P(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

Let $m := \deg(P(z))$. Then an $(nm + r) \times (nm + r)$ matrix pencil of the form

$$\mathbb{L}(z) := \left[\begin{array}{c|c} X + zY & \mathcal{C} \\ \hline \mathcal{B} & A - zI_r \end{array}\right]$$

is called a Rosenbrock linearization of $\mathbf{S}(z)$ if there exist $nm \times nm$ unimodular matrix polynomials $U(z)$ and $V(z)$ such that

$$\left[\begin{array}{c|c} U(z) & 0 \\ \hline 0 & I_r \end{array}\right] \mathbb{L}(z) \left[\begin{array}{c|c} V(z) & 0 \\ \hline 0 & I_r \end{array}\right] = \left[\begin{array}{c|c} I_{(m-1)n} & 0 \\ \hline 0 & \mathbf{S}(z) \end{array}\right].$$

The linearization in Definition 4.2.3 is in fact a linearization of $G(z)$ in the sense of Definition 4.2.2; see [2]. The advantage of Definition 4.2.3 is that it can be used to construct a linearization of $G(z)$ from a linearization of $P(z)$.

We are now ready to extend the concept of linearization of a rational matrix to the case of a meromorphic matrix when $\#(\wp_\Omega(\mathbf{M})) < \infty$. This is the case when Ω is a compact region. Let X and Y be Banach spaces. Let $T_1, T_2 \in \mathbb{H}(\Omega, \mathcal{L}(X, Y))$. Then recall that $T_1(z) \sim_\Omega T_2(z)$ if there are exist $U \in \mathbb{H}(\Omega, \mathcal{L}(Y))$ and $V \in \mathbb{H}(\Omega, \mathcal{L}(X))$ such that $U(z)$ and $V(z)$ are invertible for all $z \in \Omega$ and $T_2(z) = U(z)T_1(z)V(z)$. We denote the identity operator on X by I_X whenever it is necessary.

Definition 4.2.4. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$ *be such that* $\#(\wp_\Omega(\mathbf{M})) < \infty$. *Let* $\mathbf{M}(z) = N(z)D(z)^{-1}$ *be a right coprime MFD, where* $D(z) \in \mathbb{C}[z]^{n\times n}$ *is regular and* $N \in \mathbb{H}(\Omega)^{n\times n}$. *Set* $r := \deg(\det(D(z)))$ *and* $p := \max(r, n)$. *Let* $\mathcal{X}$ *be a Banach space and* $T \in \mathcal{L}(\mathcal{X})$. *Then a bounded operator pencil* $\mathbb{L}(z)$ *of the form*

$$\mathbb{L}(z) := \left[\begin{array}{c|c} T - zI & C_1 + zC_2 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right] : \mathcal{X} \oplus \mathbb{C}^r \to \mathcal{X} \oplus \mathbb{C}^r, \tag{4.6}$$

is said to be a linearization of $\mathbf{M}(z)$ *if the following conditions hold:*

(a) The operator $\begin{bmatrix} B_1 + zB_2 & A - zI_r \end{bmatrix} : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathbb{C}^r$ *is right invertible and the operator* $\begin{bmatrix} C_1 + zC_2 \\ A - zI_r \end{bmatrix} : \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r$ *is left invertible for all* $z \in \Omega$.

(b) There exist a Banach space $\mathcal{Z}$ *such that*

$$\left[\begin{array}{c|c} \mathbb{L}(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|cc} I_{\mathcal{Z}} & & 0 \\ \hline 0 & N(z) & 0 \\ & 0 & I_{p-n} \end{array}\right] \text{ and } \left[\begin{array}{c|c} I_{p-n} & 0 \\ \hline 0 & D(z) \end{array}\right] \sim_\Omega \left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & A - zI_r \end{array}\right].$$

We now define a linearization of the system matrix by considering a minimal local realization of $\mathbf{M}(z)$.

Definition 4.2.5. *Let* $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$. *Let* $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ *be a minimal local realization. Consider the system matrix*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

Let $\mathcal{X}$ *be a Banach space and* $T \in \mathcal{L}(\mathcal{X})$. *Then a bounded operator pencil* $\mathbb{L}(z)$ *of the form*

$$\mathbb{L}(z) := \left[\begin{array}{c|c} T - zI & C_1 + zC_2 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right] : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r$$

is said to be a linearization of $\mathbf{S}(z)$ *if the following conditions hold:*

(a) The operator $\begin{bmatrix} B_1 + zB_2 & A - zI_r \end{bmatrix} : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathbb{C}^r$ *is right invertible and the operator* $\begin{bmatrix} C_1 + zC_2 \\ A - zI_r \end{bmatrix} : \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r$ *is left invertible for all* $z \in \Omega$.

(b) *There is a Banach space $\mathcal{Z}$ and holomorphic invertible operator-valued functions $U : \Omega \longrightarrow \mathcal{L}(\mathcal{X}, \mathcal{Z} \oplus \mathbb{C}^n)$ and $V : \Omega \longrightarrow \mathcal{L}(\mathcal{Z} \oplus \mathbb{C}^n, \mathcal{X})$ such that*

$$\left[\begin{array}{c|c} U(z) & \\ \hline & I_r \end{array}\right] \mathbb{L}(z) \left[\begin{array}{c|c} V(z) & \\ \hline & I_r \end{array}\right] = \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{S}(z) \end{array}\right].$$

If $\mathbb{L}(z)$ is a linearization of $\mathbf{S}(z)$ then it turns out that $\mathbb{L}(z)$ is a linearization of $\mathbf{M}(z)$.

Theorem 4.2.6. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$ and $\mathbf{M}(z) = N(z)D(z)^{-1}$ be a right coprime MFD. Let $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ be a minimal local realization, where $A \in \mathbb{C}^{r\times r}$ and $\mathbf{H} \in \mathbb{H}(\Omega)^{n\times n}$. Consider the system matrix*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

Let $\mathbb{L}(z)$ be a linearization of $\mathbf{S}(z)$. Set $p := \max(n, r)$. Then there exists a Banach space $\mathcal{Z}$ such that

$$\left[\begin{array}{c|c} \mathbb{L}(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|cc} I_{\mathcal{Z}} & & 0 \\ \hline 0 & N(z) & 0 \\ & 0 & I_{p-n} \end{array}\right] \text{ and } \left[\begin{array}{c|c} I_{p-n} & 0 \\ \hline 0 & D(z) \end{array}\right] \sim_\Omega \left[\begin{array}{c|c} I_{p-r} & 0 \\ \hline 0 & A - zI_r \end{array}\right].$$

In other words, $\mathbb{L}(z)$ is a linearization of $\mathbf{M}(z)$.

Proof. Since $\mathbb{L}(z)$ is a linearization of $\mathbf{S}(z)$, by Definition 4.2.5, we have

$$\mathbb{L}(z) \sim_\Omega \left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(z) \end{array}\right] \Longrightarrow \left[\begin{array}{c|c} \mathbb{L}(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|cc} I_{\mathcal{Z}} & & 0 \\ \hline 0 & \mathbf{S}(z) & 0 \\ & 0 & I_{p-r} \end{array}\right]$$

Since $\mathbf{M}(z) = N(z)D(z)^{-1}$ is a coprime MFD and $\mathbf{S}(z)$ is an irreducible system matrix, by Theorem 4.1.18, we have

$$\left[\begin{array}{c|c} S(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|c} N(z) & 0 \\ \hline 0 & I_{p-n} \end{array}\right] \text{ and } \left[\begin{array}{c|c} A - zI_r & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|c} D(z) & 0 \\ \hline 0 & I_{p-n} \end{array}\right].$$

Hence it follows that

$$\left[\begin{array}{c|c} \mathbb{L}(z) & 0 \\ \hline 0 & I_{p-r} \end{array}\right] \sim_\Omega \left[\begin{array}{c|cc} I_{\mathcal{Z}} & & 0 \\ \hline 0 & N(z) & 0 \\ & 0 & I_{p-n} \end{array}\right].$$

This shows that $\mathbb{L}(z)$ is linearization of $\mathbf{M}(z)$. □

Let $\mathbb{G}(z) := T - zI + (C_1 + zC_2)(zI_r - A)^{-1}(B_1 + zB_2)$ be the transfer function of $\mathbb{L}(z)$ given in Definition 4.2.5. Then the next result shows that $\mathbb{G}(z) \sim_\Omega I_{\mathcal{Z}} \oplus \mathbf{M}(z)$.

Theorem 4.2.7. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$. Let $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ be a minimal local realization, where $A \in \mathbb{C}^{r\times r}$ and $\mathbf{H} \in \mathbb{H}(\Omega)^{n\times n}$. Consider the system matrix*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

Let $\mathcal{X}$ be a Banach space and $\mathbb{L}(z)$ be a linearization of $\mathbf{S}(z)$ given by

$$\mathbb{L}(z) := \left[\begin{array}{c|c} T - zI & C_1 + zC_2 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right] : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r.$$

Let $\mathcal{Z}$ be a Banach space and $U : \Omega \longrightarrow \mathcal{L}(\mathcal{X},\, \mathcal{Z} \oplus \mathbb{C}^n)$ and $V : \Omega \longrightarrow \mathcal{L}(\mathcal{Z} \oplus \mathbb{C}^n,\, \mathcal{X})$ be holomorphic invertible operator-valued functions such that

$$\left[\begin{array}{c|c} U(z) & \\ \hline & I_r \end{array}\right] \mathbb{L}(z) \left[\begin{array}{c|c} V(z) & \\ \hline & I_r \end{array}\right] = \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{S}(z) \end{array}\right]. \tag{4.7}$$

Let $\mathbb{G}(z) := T - zI + (C_1 + zC_2)(zI_r - A)^{-1}(B_1 + zB_2)$ be the transfer function of $\mathbb{L}(z)$. Then we have

$$U(z)\mathbb{G}(z)V(z) = \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{M}(z) \end{array}\right] \quad \text{and} \quad \sigma_\Omega(\mathbb{G}) = \sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{S}) = \sigma_\Omega(\mathbb{L}).$$

Proof. Equating the blocks in (4.7), we have

$$U(z)(T - zI)V(z) = \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{H}(z) \end{array}\right],\ U(z)(C_1 + zC_2) = \begin{bmatrix} 0 \\ C \end{bmatrix},\ (B_1 + zB_2)V(z) = \begin{bmatrix} 0 & B \end{bmatrix}.$$

Hence

$$\begin{aligned} U(z)\mathbb{G}(z)V(z) &= U(z)(T - zI)V(z) + U(z)(C_1 + zC_2)(zI_r - A)^{-1}(B_1 + zB_2)V(z) \\ &= \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{H}(z) \end{array}\right] + \begin{bmatrix} 0 \\ C \end{bmatrix} (zI_r - A)^{-1} \begin{bmatrix} 0 & B \end{bmatrix} \\ &= \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{H}(z) + C(zI_r - A)^{-1}B \end{array}\right] = \left[\begin{array}{c|c} I_{\mathcal{Z}} & \\ \hline & \mathbf{M}(z) \end{array}\right]. \end{aligned}$$

This shows that $\mathbb{G}(z) \sim_\Omega I_{\mathcal{Z}} \oplus \mathbf{M}(z) \Longrightarrow \sigma_\Omega(\mathbb{G}) = \sigma_\Omega(\mathbf{M})$. Evidently, $\sigma_\Omega(\mathbb{L}) = \sigma_\Omega(\mathbf{S})$. As $\mathbf{S}(z)$ is irreducible, we have $\sigma_\Omega(\mathbb{L}) = \sigma_\Omega(\mathbf{S}) = \sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbb{G})$. $\square$

The proof of Theorem 4.2.7 shows that if $\mathbb{L}(z)$ is a linearization of $\mathbf{S}(z)$ then $T-zI \sim_\Omega I_\mathcal{Z} \oplus \mathbf{H}(z)$, that is, $T-zI$ is a linearization of $\mathbf{H}(z)$. Therefore, the question that arises is this: If $T-zI$ is a linearization of $\mathbf{H}(z)$ then is it possible to construct a linearization $\mathbb{L}(z)$ of $\mathbf{S}(z)$? The next result provides an answer.

Theorem 4.2.8. *Let $\Omega \subset \mathbb{C}$ be a bounded simply connected domain such that the boundary $\partial\Omega$ is a positively oriented rectifiable simple closed curve. Let $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$. Let $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ be a minimal local realization. Consider*

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right].$$

Let $z_0 \in \Omega$. Define $B_1, B_2 : \mathcal{C}(\partial\Omega, \mathbb{C}^n) \longrightarrow \mathbb{C}^r$ by

$$B_1 f := -\frac{1}{2\pi i}\int_{\partial\Omega} \frac{wBf(w)}{w - z_0}dw \quad and \quad B_2 f := \frac{1}{2\pi i}\int_{\partial\Omega}\frac{Bf(w)}{w-z_0}dw$$

Define $C_1 : \mathbb{C}^r \longrightarrow \mathcal{C}(\partial\Omega,\ \mathbb{C}^n), u \longmapsto C_1u,$ by $(C_1u)(z) := Cu$ for all $z \in \Omega$. Suppose that $\mathbf{H}(z)$ is regular on $\partial\Omega$. Let T be the Gohberg-Kaashoek-Lay linearization of $\mathbf{H}(z)$ as given in Theorem 3.4.3. Then

$$\mathbb{L}(z) := \left[\begin{array}{c|c} T - zI & C_1 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right] = \left[\begin{array}{c|c} T & C_1 \\ \hline B_1 & A \end{array}\right] - z\left[\begin{array}{c|c} I & 0 \\ \hline -B_2 & I_r \end{array}\right]$$

is a linearization of $\mathbf{S}(z)$. Further, we have $\sigma_\Omega(\mathbf{M}) = \sigma_\Omega(\mathbf{S}) = \sigma_\Omega(\mathbb{L})$ and $\partial\Omega \subset \sigma_{\mathrm{ap}}(\mathbb{L})$, where $\sigma_{\mathrm{ap}}(\mathbb{L})$ is the approximate spectrum of $\mathbb{L}$.

Proof. Set $\mathcal{X} := \mathcal{C}(\partial\Omega, \mathbb{C}^n)$. For $f \in \mathcal{X}$ and any norm $\|\cdot\|$ on $\mathbb{C}^n$, consider the norm on $\mathcal{X}$ given by

$$\|f\|_\infty := \sup\{\|f(z)\| : z \in \partial\Omega\}.$$

Then $\mathcal{X}$ is a Banach space. The sup norm on $\mathcal{X} \oplus \mathbb{C}^r$ is defined similarly. Consider the Banach space

$$\mathcal{Z} := \left\{ f \in \mathcal{C}(\partial\Omega, \mathbb{C}^n) : \int_{\partial\Omega} \frac{f(\omega)}{\omega - z_0} d\omega = 0 \right\}.$$

As shown in the proof of Theorem 3.4.3, there exist holomorphic invertible operator-valued functions $F(z) : \mathcal{X} \to \mathcal{Z} \oplus \mathbb{C}^n$ and $E(z) : \mathcal{Z} \oplus \mathbb{C}^n \to \mathcal{X}$ such that

$$F(z)(zI - T)E(z) = \left[\begin{array}{c|c} I_\mathcal{Z} & 0 \\ \hline 0 & \mathbf{H}(z) \end{array}\right] \text{ for all } z \in \Omega. \tag{4.8}$$

We have

$$\left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(z) \end{array}\right] = \left[\begin{array}{c|cc} I_{\mathcal{Z}} & 0 & 0 \\ \hline 0 & \mathbf{H}(z) & C \\ 0 & B & A - zI_r \end{array}\right] = \left[\begin{array}{cc|c} I_{\mathcal{Z}} & 0 & 0 \\ 0 & \mathbf{H}(z) & C \\ \hline 0 & B & A - zI_r \end{array}\right]$$

$$= \left[\begin{array}{c|c} F(z)(zI-T)E(z) & \begin{matrix} 0 \\ C \end{matrix} \\ \hline \begin{matrix} 0 & B \end{matrix} & A - zI_r \end{array}\right]$$

$$= \left[\begin{array}{c|c} F(z) & 0 \\ \hline 0 & I_r \end{array}\right] \left[\begin{array}{c|c} zI - T & F(z)^{-1}\begin{bmatrix} 0 \\ C \end{bmatrix} \\ \hline \begin{bmatrix} 0 & B \end{bmatrix} E(z)^{-1} & A - zI_r \end{array}\right] \left[\begin{array}{c|c} E(z) & 0 \\ \hline 0 & I_r \end{array}\right].$$

We need to show that $F(z)^{-1}\begin{bmatrix} 0 \\ C \end{bmatrix} = C_1$ and $\begin{bmatrix} 0 & B \end{bmatrix} E(z)^{-1} = B_1 + zB_2$.

The operators $E(z)$ and $F(z)$ in (4.8) are constructed in the proof of Theorem 3.4.3 in Chapter 3. For convenience, we briefly mention these operators. For $z \in \Omega$, $E(z) : Z \oplus \mathbb{C}^n \longrightarrow \mathcal{X}$ is given by

$$E(z) := (zI - V)^{-1} J \tag{4.9}$$

and $F(z) : \mathcal{X} \longrightarrow \mathcal{Z} \oplus \mathbb{C}^n$ is given by

$$F(z) := J^{-1}(I - PB(z)(I - P)). \tag{4.10}$$

The multiplication operator $V : \mathcal{X} \to \mathcal{X}$, $f \longmapsto Vf$, is given by

$$(Vf)(w) = wf(w), \ \forall\, w \in \partial\Omega. \tag{4.11}$$

For $z \in \Omega$, the operator $zI - V$ is invertible and $(zI - V)^{-1} f(w) = (z - w)^{-1} f(w)$ for $w \in \partial\Omega$ and $f \in \mathcal{X}$. The operator $B(z)$ is defined in the proof of Theorem 3.4.3. The operators P and J are defined as follows.

Consider the canonical embedding $\mathcal{I} : \mathbb{C}^n \to \mathcal{X}$ defined by $(\mathcal{I}u)(z) := u$ for $z \in \partial\Omega$. Define $\mathcal{W} : \mathcal{X} \to \mathbb{C}^n$ by

$$\mathcal{W}(f) = \frac{1}{2\pi i} \int_{\partial\Omega} \frac{f(w)}{w - z_0} dw.$$

Then as shown in the proof of Theorem 3.4.3, $\mathcal{W}\mathcal{I} = I_n$ and $P := \mathcal{I}\mathcal{W} : \mathcal{X} \to \mathcal{X}$ is a projection. Further, $\mathcal{Z}$ is the null space of $\mathcal{W}$. Consider $J : \mathcal{Z} \oplus \mathbb{C}^n \to \mathcal{X}$ given by

$$J\left(\begin{bmatrix} g \\ u \end{bmatrix}\right) = g + \mathcal{I}u,\ \forall \begin{bmatrix} g \\ u \end{bmatrix} \in \mathcal{Z} \oplus \mathbb{C}^n.$$

Then J is invertible and $J^{-1}(f) = \begin{bmatrix} (I-P)f \\ \mathcal{W}f \end{bmatrix}$ for all $f \in \mathcal{X}$. Now

$$\begin{bmatrix} 0 & B \end{bmatrix} E(z)^{-1} = \begin{bmatrix} 0 & B \end{bmatrix} J^{-1}(zI - V) = \begin{bmatrix} 0 & B \end{bmatrix} \begin{bmatrix} (I-P) \\ \mathcal{W}(zI-V) \end{bmatrix} \tag{4.12}$$

$$= B\mathcal{W}(zI - V) = zB\mathcal{W} - B\mathcal{W}V = zB_2 + B_1 \tag{4.13}$$

where $B_2 := B\mathcal{W} : \mathcal{X} \to \mathbb{C}^r$ defined by

$$B_2 f = (B\mathcal{W})f = B\left(\frac{1}{2\pi i}\int_{\partial\Omega} \frac{f(w)}{w - z_0}dw\right)$$

and $B_1 := B\mathcal{W}V : \mathcal{X} \to \mathbb{C}^r$ by

$$B_1 f = (B\mathcal{W}V)\, f = -B\left(\frac{1}{2\pi i}\int_{\partial\Omega} \frac{wf(w)}{w - z_0}dw\right) \quad \text{for all } f \in \mathcal{C}(\partial\Omega, \mathbb{C}^n).$$

Now $F(z) = J^{-1}(I - PB(z)(I-P)) \Longrightarrow F(z)^{-1} = (I + PB(z)(I-P))J$. Hence

$$\begin{aligned} F(z)^{-1}\begin{bmatrix} 0 \\ C \end{bmatrix} &= (I + PB(z)(I-P))J\begin{bmatrix} 0 \\ C \end{bmatrix} \\ &= (I + PB(z)(I-P))\mathcal{I}C \\ &= \mathcal{I}C + PB(z)(I-P))\mathcal{I}C. \end{aligned}$$

Since $P = \mathcal{I}\mathcal{W}$ and $\mathcal{W}\mathcal{I} = I_n$, we have $(I-P)\mathcal{I}C = (\mathcal{I}C - \mathcal{I}\mathcal{W}\mathcal{I}C) = \mathcal{I}C - \mathcal{I}C = 0$. This shows that $F(z)^{-1}\begin{bmatrix} 0 \\ C \end{bmatrix} = \mathcal{I}C = C_1$. We show that $\mathbb{L}(z)$ is irreducible, that is,

$$\begin{bmatrix} B_1 + zB_2 & A - z_r I \end{bmatrix} : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathbb{C}^r$$

is right invertible and

$$\begin{bmatrix} C_1 \\ A - zI_r \end{bmatrix} : \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r$$

is left invertible. Since the system matrix $\mathbf{S}(z)$ is irreducible, there exist $X \in \mathbb{C}^{n\times r}$ and $Y \in \mathbb{C}^{r\times r}$ such that $\begin{bmatrix} B & A - zI_r \end{bmatrix} \left[\begin{array}{c} X \\ \hline Y \end{array}\right] = I_r$. We now show that the operator $\left[\begin{array}{c} E(z)\begin{bmatrix} 0 \\ X \end{bmatrix} \\ \hline Y \end{array}\right] : \mathbb{C}^r \longrightarrow \mathcal{X} \oplus \mathbb{C}^r$ is a right inverse of $\begin{bmatrix} B_1 + zB_2 & A - zI_r \end{bmatrix}$. Indeed, we have (recall that $E(z) : Z \oplus \mathbb{C}^n \to \mathcal{X}$)

$$\begin{aligned}
&\begin{bmatrix} B_1 + zB_2 & A - zI_r \end{bmatrix} \begin{bmatrix} E(z)\begin{bmatrix} 0 \\ X \end{bmatrix} \\ Y \end{bmatrix} \\
&= \begin{bmatrix} [0_{\mathcal{Z}\to\mathbb{C}^r} \;\; B]\, E(z)^{-1} & A - zI_r \end{bmatrix} \begin{bmatrix} E(z)\begin{bmatrix} 0 \\ X \end{bmatrix} \\ Y \end{bmatrix} \\
&= [0_{\mathcal{Z}\to\mathbb{C}^r} \;\; B]\, E(z)^{-1}E(z) \begin{bmatrix} 0 \\ X \end{bmatrix} + (A - zI_r)Y \\
&= [0_{\mathcal{Z}\to\mathbb{C}^r} \;\; B]\, I_{Z\oplus\mathbb{C}^n} \begin{bmatrix} 0 \\ X \end{bmatrix} + (A - zI_r)Y \\
&= 0_{r\times r} + BX + (A - zI_r)Y = I_r.
\end{aligned}$$

This shows that $\begin{bmatrix} B_1 + zB_2 & A - zI_r \end{bmatrix}$ is right invertible.

Again, since $\mathbf{S}(z)$ is irreducible, there exist $X \in \mathbb{C}^{r\times n}$ and $Y \in \mathbb{C}^{r\times r}$ such that

$$\begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} C \\ A - zI_r \end{bmatrix} = I_r.$$

We show that the operator $\begin{bmatrix} [0_{\mathcal{Z}\to\mathbb{C}^r} \;\; X]\, F(z) & Y \end{bmatrix} : \mathcal{X} \oplus \mathbb{C}^r \longrightarrow \mathbb{C}^r$ is a left inverse of $\begin{bmatrix} C_1 \\ A - zI_r \end{bmatrix}$. We have (recall that $F(z) : \mathcal{X} \to Z \oplus \mathbb{C}^n$)

$$\begin{aligned}
&\begin{bmatrix} [0_{\mathcal{Z}\to\mathbb{C}^r} \quad X]\, F(z) & Y \end{bmatrix} \begin{bmatrix} C_1 \\ A - zI_r \end{bmatrix} \\
&= \begin{bmatrix} [0_{\mathcal{Z}\to\mathbb{C}^r} \quad X]\, F(z) & Y \end{bmatrix} \begin{bmatrix} F(z)^{-1} \begin{bmatrix} 0_{\mathbb{C}^r\to Z} \\ C \end{bmatrix} \\ A - zI_r \end{bmatrix} \\
&= XC + Y(A - zI_r) = I_r.
\end{aligned}$$

This shows that $\begin{bmatrix} C_1 \\ A - zI_r \end{bmatrix}$ is left invertible. Hence $\mathbb{L}(z) = \left[\begin{array}{c|c} T - zI & C_1 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right]$ is irreducible. This proves that $\mathbb{L}(z)$ is a linearization of the system matrix $\mathbf{S}(z)$.

Since $\mathbb{L}(z)$ is a linearization of $\mathbf{S}(z)$, by Theorem 4.2.7, we have $\sigma_\Omega(\mathbb{L}) = \sigma_\Omega(\mathbf{S})$. Now, we show that and $\partial\Omega \subset \sigma_{\text{ap}}(\mathbb{L})$, that is, if $\lambda_0 \in \partial\Omega$ then there exists a sequence $(\psi_k) \in \mathcal{X} \oplus \mathbb{C}^r$ such that $\|\psi_k\|_\infty = 1$ and $\|\mathbb{L}(\lambda_0)\psi_k\|_\infty \longrightarrow 0$ as $k \to \infty$.

Our construction of (ψ_k) is inspired by a procedure outlined in [27, p.116]. Let $\lambda_0 \in \partial\Omega$. Choose a sequence $(\phi_k) \subset \mathcal{C}(\partial\Omega, \mathbb{R})$ such that for all k we have $0 \le \phi_k(\lambda_0) \le 1$,

$$\phi_k(\lambda_0) = 1 \text{ and } \phi_k(z) = 0, \text{ if } |z - \lambda_0| \ge \frac{1}{k}.$$

For example, such a sequence is given by

$$\phi_k(z) := \begin{cases} 1 - k|z - \lambda_0| & \text{if } |z - \lambda_0| < \frac{1}{k} \\ 0 & \text{if } |z - \lambda_0| \ge \frac{1}{k} \end{cases}.$$

Let e_1 be the first column of the identity matrix I_n. Then $(\phi_k e_1) \subset \mathcal{X}$. Now define

$$\psi_k := \begin{bmatrix} \phi_k e_1 \\ 0 \end{bmatrix} \in \mathcal{X} \oplus \mathbb{C}^r \Longrightarrow \|\psi_k\|_\infty = 1 \quad \text{for} \quad k \in \mathbb{N}.$$

We need to show that $\|\mathbb{L}(\lambda_0)\psi_k\|_\infty = \left\| \left[\frac{(T - \lambda_0 I)\phi_k e_1}{(B_1 + \lambda_0 B_2)\phi_k e_1} \right] \right\|_\infty \longrightarrow 0.$

We have

$$\|\mathbb{L}(\lambda_0)\psi_k\|_\infty = \left\| \begin{bmatrix} (T-\lambda_0 I)\phi_k e_1 \\ (B_1+\lambda_0 B_2)\phi_k e_1 \end{bmatrix} \right\|_\infty$$
$$\leq \|(T-\lambda_0 I)\phi_k e_1\|_\infty + \|(B_1+\lambda_0 B_2)\phi_k e_1\|_\infty .$$

It is shown in [27, Theorem 2.3, p.116] that $\|(T-\lambda_0 I)\phi_k e_1\|_\infty \longrightarrow 0$ as $k \to \infty$. By (4.13), we have

$$\|(B_1+\lambda_0 B_2)\phi_k e_1\|_\infty = \| B\mathcal{W}(\lambda_0 I - V)\phi_k e_1\|_\infty$$
$$\leq \|B\mathcal{W}\| \, \|(\lambda_0 I - V)\phi_k e_1\|_\infty .$$

It follows from (4.11) that

$$\|(\lambda_0 I - V)\phi_k e_1\|_\infty = \sup_{z\in\partial\Omega} \|(\lambda_0 I - V)\phi_k(z) e_1\|$$
$$= \sup_{z\in\partial\Omega} \|(z-\lambda_0)\phi_k(z)\| \leq \frac{1}{k} \longrightarrow 0 \text{ as } k \to \infty.$$

This shows that $\|(B_1+\lambda_0 B_2)\phi_k e_1\|_\infty \to 0$ as $k \to \infty$. Hence $\|\mathbb{L}(\lambda_0)\psi_k\|_\infty \to 0$ as $k \to \infty$. This shows that $\lambda_0 \in \sigma_{\text{ap}}(\mathbb{L}) \Longrightarrow \partial\Omega \subset \sigma_{\text{ap}}(\mathbb{L})$. □

4.2.1 Recovery of eigenvectors from a linearization

Recall that $\mathcal{X} := \mathcal{C}(\partial\Omega,\ \mathbb{C}^n)$. Let $\mathbb{L}(z) := \left[\begin{array}{c|c} T - zI & C_1 \\ \hline B_1 + zB_2 & A - zI_r \end{array}\right] : \mathcal{X} \oplus \mathbb{C}^r \to \mathcal{X} \oplus \mathbb{C}^r$ be the linearization given in Theorem 4.2.8 of the irreducible system matrix

$$\mathbf{S}(z) := \left[\begin{array}{c|c} \mathbf{H}(z) & C \\ \hline B & A - zI_r \end{array}\right]_{(n+r)\times(n+r)} .$$

We now describe recovery of an eigenvector of $\mathbf{S}(z)$ from an eigenvector of $\mathbb{L}(z)$.

Theorem 4.2.9. *Let $\lambda \in \sigma_\Omega(\mathbf{S})$. Define $\mathcal{E}_\lambda : \mathbb{C}^n \longrightarrow \mathcal{X}$ by $(\mathcal{E}_\lambda u)(\omega) := \frac{u}{\lambda-\omega}$* for $\omega \in \partial\Omega$. *Let $z_0 \in \Omega$ be fixed. Then the following maps are linear isomorphisms*

$$\mathbb{E} : \mathcal{N}(\mathbf{S}(\lambda)) \to \mathcal{N}(\mathbb{L}(\lambda)), \qquad \left[\begin{array}{c} u \\ \hline v \end{array}\right] \longmapsto \left[\begin{array}{c} \mathcal{E}_\lambda u \\ \hline v \end{array}\right],$$

$$\mathbb{F} : \mathcal{N}(\mathbb{L}(\lambda)) \to \mathcal{N}(\mathbf{S}(\lambda)), \qquad \left[\begin{array}{c} f \\ \hline v \end{array}\right] \longmapsto \left[\begin{array}{c} \frac{1}{2\pi i}\int_\Omega \frac{\lambda-\omega}{\omega - z_0} f(w)dw \\ \hline v \end{array}\right],$$

where $u \in \mathbb{C}^n, v \in \mathbb{C}^r$ and $f \in \mathcal{X}$.

Proof. For $F(z)$ and $E(z)$ defined in (4.10) and (4.9), respectively, we have

$$\left[\begin{array}{c|c} F(z) & 0 \\ \hline 0 & I_r \end{array}\right] \mathbb{L}(z) \left[\begin{array}{c|c} E(z) & 0 \\ \hline 0 & I_r \end{array}\right] = \left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(z) \end{array}\right].$$

Note that $\begin{bmatrix} g \\ \hline u \\ v \end{bmatrix} \in \mathcal{N}\left(\left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(\lambda) \end{array}\right]\right) \iff \begin{bmatrix} g \\ \hline \mathbf{S}(\lambda)\begin{bmatrix} u \\ v \end{bmatrix} \end{bmatrix} = 0$

$$\iff g = 0 \text{ and } \begin{bmatrix} u \\ v \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda))$$

Hence it follows that $\Pi : \mathcal{N}\left(\left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(\lambda) \end{array}\right]\right) \longrightarrow \mathcal{N}(\mathbf{S}(\lambda)),\ \begin{bmatrix} 0 \\ \hline u \\ v \end{bmatrix} \longmapsto \begin{bmatrix} u \\ v \end{bmatrix}$ is well defined and is an isomorphism. Now,

$$\begin{bmatrix} u \\ v \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda)) \iff \left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(\lambda) \end{array}\right] \begin{bmatrix} 0 \\ \hline u \\ v \end{bmatrix} = 0 \iff \mathbb{L}(\lambda) \left[\begin{array}{c|c} E(\lambda) & 0 \\ \hline 0 & I_r \end{array}\right] \begin{bmatrix} 0 \\ u \\ \hline v \end{bmatrix} = 0$$

$$\iff \begin{bmatrix} E(\lambda)\begin{bmatrix} 0 \\ u \end{bmatrix} \\ v \end{bmatrix} \in \mathcal{N}\left(\mathbb{L}(\lambda)\right).$$

It follows that $E(\lambda)\begin{bmatrix} 0 \\ u \end{bmatrix} = \mathcal{E}_\lambda u$, where $(\mathcal{E}_\lambda u)(w) = \frac{u}{\lambda - \omega}$ for all $\omega \in \partial\Omega$. Thus for $\begin{bmatrix} u \\ v \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda))$, we have $\mathbb{E}\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) = \begin{bmatrix} E(\lambda)\begin{bmatrix} 0 \\ u \end{bmatrix} \\ v \end{bmatrix} = \begin{bmatrix} \mathcal{E}u \\ v \end{bmatrix}$.

$$\text{Similarly } \begin{bmatrix} f \\ v \end{bmatrix} \in \mathcal{N}(\mathbb{L}(\lambda)) \iff \left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(\lambda) \end{array}\right] \left[\begin{array}{c|c} E(\lambda)^{-1} & 0 \\ \hline 0 & I_r \end{array}\right] \begin{bmatrix} f \\ v \end{bmatrix} = 0 \quad (4.14)$$

$$\iff \begin{bmatrix} E(\lambda)^{-1}f \\ v \end{bmatrix} \in \mathcal{N}\left(\left[\begin{array}{c|c} I_{\mathcal{Z}} & 0 \\ \hline 0 & \mathbf{S}(\lambda) \end{array}\right]\right) \iff \Pi \begin{bmatrix} E(\lambda)^{-1}f \\ v \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda)). \quad (4.15)$$

Define $\mathbb{F} : \mathcal{N}(\mathbb{L}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{S}(\lambda))$ by

$$\begin{bmatrix} f \\ v \end{bmatrix} \mapsto \Pi \begin{bmatrix} E(\lambda)^{-1}f \\ v \end{bmatrix}.$$

It follows from (4.15) that $\mathbb{F}$ is well defined. For $\begin{bmatrix} f \\ v \end{bmatrix} \in \mathcal{N}(\mathbb{L}(\lambda))$ we have

$$\mathbb{F}\left(\begin{bmatrix} f \\ v \end{bmatrix}\right) = \Pi \begin{bmatrix} E^{-1}(\lambda)f \\ v \end{bmatrix} = \Pi \begin{bmatrix} J^{-1}(\lambda - V)f \\ v \end{bmatrix} = \Pi \left[\begin{array}{c} (I-P)f \\ \hline \mathcal{W}(\lambda I - V)f \\ v \end{array}\right]$$

$$= \left[\begin{array}{c} \mathcal{W}(\lambda I - V)f \\ \hline v \end{array}\right] = \left[\begin{array}{c} \frac{1}{2\pi i}\int_{\partial\Omega} \frac{\lambda - \omega}{\omega - z_0} f(w)dw \\ \hline v \end{array}\right].$$

Note that $\mathbb{F}\mathbb{E} : \mathcal{N}(\mathbf{S}(\lambda)) \to \mathcal{N}(\mathbf{S}(\lambda))$. Let $\left[\begin{array}{c} u \\ \hline v \end{array}\right] \in \mathcal{N}(\mathbf{S}(\lambda))$. Then

$$\mathbb{F}\mathbb{E}\left[\begin{array}{c} u \\ \hline v \end{array}\right] = \mathbb{F}\left[\begin{array}{c|c} E(z) & 0 \\ \hline 0 & I_r \end{array}\right]\left[\begin{array}{c} 0 \\ u \\ \hline v \end{array}\right] = \mathbb{F}\left[\begin{array}{c} E(z)\begin{bmatrix} 0 \\ u \end{bmatrix} \\ \hline v \end{array}\right]$$

$$= \Pi \begin{bmatrix} E(z)^{-1}E(z)\begin{bmatrix} 0 \\ u \end{bmatrix} \\ v \end{bmatrix} = \Pi \left[\begin{array}{c} 0 \\ \hline u \\ v \end{array}\right] = \begin{bmatrix} u \\ v \end{bmatrix}$$

which shows that $\mathbb{F}\mathbb{E} = I_{n+r}$. Hence it follows that $\mathbb{E}$ is one-one and $\mathbb{F}$ is onto.

Let $\begin{bmatrix} f \\ \hline v \end{bmatrix} \in \mathcal{N}(\mathbb{L}(\lambda))$ and $\mathbb{F}\left(\begin{bmatrix} f \\ \hline v \end{bmatrix}\right) = 0$. Then

$$\mathbb{F}\left(\begin{bmatrix} f \\ \hline v \end{bmatrix}\right) = \Pi \begin{bmatrix} E(\lambda)^{-1}f \\ \hline v \end{bmatrix} = 0 \implies \begin{bmatrix} E(\lambda)^{-1}f \\ \hline v \end{bmatrix} = 0 \implies \begin{bmatrix} f \\ \hline v \end{bmatrix} = 0$$

which shows that $\mathbb{F}$ is one-one. Consequently, the maps $\mathbb{E} : \mathcal{N}(\mathbf{S}(\lambda)) \to \mathcal{N}(\mathbb{L}(\lambda))$ and $\mathbb{F} : \mathcal{N}(\mathbb{L}(\lambda)) \to \mathcal{N}(\mathbf{S}(\lambda))$ are linear isomorphisms. □

The next result describes the recovery of eigenvectors $\mathbf{M}(z)$ from those of a linearization of $\mathbf{M}(z)$.

Theorem 4.2.10. *Let $\mathbf{M}(z) = \mathbf{H}(z) + C(zI_r - A)^{-1}B$ be a minimal local realization. Let $\mathbb{L}(z)$ be the linearization given in Theorem 4.2.8. Let $\lambda \in \sigma_\Omega(\mathbf{M})$. Let $z_0 \in \Omega$ be fixed. Define $\mathtt{E}_\lambda : \mathcal{N}(\mathbf{M}(\lambda)) \longrightarrow \mathcal{N}(\mathbb{L}(\lambda))$ and $\mathtt{F}_\lambda : \mathcal{N}(\mathbb{L}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{M}(\lambda))$ by*

$$\mathtt{E}_\lambda x := \begin{bmatrix} \mathcal{E}_\lambda x \\ (\lambda I_r - A)^{-1}Bx \end{bmatrix} \text{ and } \mathtt{F}_\lambda\left(\begin{bmatrix} f \\ y \end{bmatrix}\right) = \frac{1}{2\pi i}\int_{\partial\Omega} \frac{(\lambda-\omega)f(\omega)}{z_0-\omega}d\omega,$$

where $(\mathcal{E}_\lambda x)(\omega) := \frac{x}{\lambda-\omega}$ for $\omega \in \partial\Omega$. Then $\mathtt{E}_\lambda$ and $\mathtt{F}_\lambda$ are linear isomorphisms.

Proof. Define the maps $f : \mathcal{N}(\mathbf{M}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{S}(\lambda))$ and $g : \mathcal{N}(\mathbf{S}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{M}(\lambda))$ by

$$f(x) = \begin{bmatrix} x \\ (\lambda I_r - A)^{-1}Bx \end{bmatrix} \text{ and } g\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x.$$

Let $\mathbf{M}(\lambda)x = 0$. Then

$$\mathbf{S}(\lambda)\begin{bmatrix} x \\ (\lambda I_r - A)^{-1}Bx \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{H}(\lambda) & C \\ \hline B & A-\lambda I_r \end{array}\right]\begin{bmatrix} x \\ (\lambda I_r - A)^{-1}Bx \end{bmatrix} = \begin{bmatrix} \mathbf{M}(\lambda)x \\ Bx - Bx \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which shows that $f(x) \in \mathcal{N}(\mathbf{S}(\lambda))$ and f is well defined. Obviously f is injective. On the other hand, if $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda))$ then it is easy to see that $\mathbf{M}(\lambda)x = 0$ and $y = (\lambda I_r - A)^{-1}Bx$. Hence $\begin{bmatrix} x \\ y \end{bmatrix} = f(x)$ and $x \in \mathcal{N}(\mathbf{M}(\lambda))$. This shows that $f : \mathcal{N}(\mathbf{M}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{S}(\lambda))$ is a linear isomorphism.

Next, let $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda))$ and $g\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = 0$. Then

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda)) \implies \left[\begin{array}{c|c} \mathbf{H}(\lambda) & C \\ \hline B & A - \lambda I_r \end{array}\right] \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Longrightarrow y = -(\lambda I_r - A)^{-1} B x$$

which shows that $x = 0 \implies y = 0 \Longrightarrow g$ is one-one. On the other hand,

$$x \in \mathcal{N}(\mathbf{M}(\lambda)) \implies \begin{bmatrix} x \\ -(A - \lambda I_r) B x \end{bmatrix} \in \mathcal{N}(\mathbf{S}(\lambda)).$$

This shows that g is onto. Hence g is a linear isomorphism.

Since $f : \mathcal{N}(\mathbf{M}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{S}(\lambda))$ and $g : \mathcal{N}(\mathbf{S}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{M}(\lambda))$ are isomorphism and, by Theorem 4.2.9, $\mathbb{E} : \mathcal{N}(\mathbf{S}(\lambda)) \to \mathcal{N}(\mathbb{L}(\lambda))$ and $\mathbb{F} : \mathcal{N}(\mathbb{L}(\lambda)) \to \mathcal{N}(\mathbf{S}(\lambda))$ are isomorphisms, it follows that

$$\mathtt{E}_\lambda = \mathbb{E} o f : \mathcal{N}(\mathbf{M}(\lambda)) \to \mathcal{N}(\mathbb{L}(\lambda)) \text{ and } \mathtt{F}_\lambda = g o \mathbb{F} : \mathcal{N}(\mathbb{L}(\lambda)) \longrightarrow \mathcal{N}(\mathbf{M}(\lambda))$$

are linear isomorphisms. This completes the proof. □

5

Moment problems for eigenvalues and spectral projections

Let A be a bounded linear operator with nonempty discrete spectrum $\sigma_d(A)$. Let Ω be a bounded domain in the complex plane $\mathbb{C}$ with rectifiable boundary $\Gamma := \partial\Omega$ such that $\Gamma \subset \rho(A)$ and $\sigma(A) \cap \Omega = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$, where $\rho(A)$ and $\sigma(A)$ are the resolvent set and the spectrum of A, respectively. We consider the *tracial moments*

$$s_p := \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma z^p(zI - A)^{-1}dz\right),\ p \in \mathbb{N} \cup \{0\}.$$

Given the moments $s_0, \ldots, s_p$, for an appropriate $p \in \mathbb{N}$, the task is to recover the ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A and their algebraic multiplicities. We show that exactly 2ℓ tracial moments, namely, $s_0, \ldots, s_{2\ell-1}$ are needed to recover the eigenvalues $\lambda_1, \ldots, \lambda_\ell$. Next, we consider the *operator moments*

$$S_p := \frac{1}{2\pi i}\int_\Gamma z^p(zI - A)^{-1}dz,\ p \in \mathbb{N} \cup \{0\}.$$

Given the moments $S_0, \ldots, S_p$, for an appropriate $p \in \mathbb{N}$, the task is to recover the spectral projections $P_1, \ldots, P_\ell$ of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively. If the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ are semisimple, then we show that exactly ℓ operator moments, namely, $S_0, \ldots, S_{\ell-1}$ are needed to recover the spectral projections $P_1, \ldots, P_\ell$.

5.1 Introduction

Let X be a complex Banach space and $A : X \longrightarrow X$ be a bounded linear operator. Let $\sigma(A), \sigma_d(A)$ and $\rho(A)$, respectively, denote the spectrum, discrete spectrum and the resolvent set of A. We assume that the discrete spectrum $\sigma_d(A)$ is nonempty. Let

$\Omega \subset \mathbb{C}$ be a bounded domain such that the boundary $\Gamma := \partial\Omega$ is rectifable, simple, closed and positively oriented, $\Gamma \subset \rho(A)$ and that $\sigma_0 := \sigma(A) \cap \Omega = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. Thus Ω contains ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A of finite algebraic multiplicity. Let $P_1, \ldots, P_\ell$, respectively, be the spectral projections of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$. Now, consider the *tracial moments*

$$s_p := \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma z^p(zI - A)^{-1}dz\right),\ p \in \mathbb{N}\cup\{0\},$$

where $\mathrm{Tr}(B)$ denotes the trace of a trace-class operator B. Then the moment problem for the eigenvalues of A can be stated as follows.

Problem-A: *Given the moments $s_0, \ldots, s_p$, for an appropriate $p \in \mathbb{N}$, determine the ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A and their algebraic multiplicities.*

We show that exactly 2ℓ moments, namely, $s_0, \ldots, s_{2\ell-1}$ are needed in order to determine the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ and their algebraic multiplicities. In fact, $\lambda_1, \ldots, \lambda_\ell$ are eigenvalues of an $\ell\times\ell$ Hankel matrix pencil $\widehat{H}_\ell - \lambda H_\ell$ irrespective of the multiplicities of the eigenvalues, where H_ℓ is nonsingular.

Next, consider the *operator moments*

$$S_p := \frac{1}{2\pi i}\int_\Gamma z^p(zI - A)^{-1}dz,\ p \in \mathbb{N}\cup\{0\}.$$

The moment problem for the spectral projections is as follows.

Problem-B: *Given the moments $S_0, \ldots, S_p$, for an appropriate $p \in \mathbb{N}$, determine the spectral projections $P_1, \ldots, P_\ell$ of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively.*

If the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ are semisimple, then we show that exactly ℓ operator moments, namely, $S_0, \ldots, S_{\ell-1}$ are needed in order to recover the spectral projections $P_1, \ldots, P_\ell$. Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ eigenpairs of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. Let the eigenvector v_j be given by $v_j := \begin{bmatrix}\beta_1^j & \cdots & \beta_\ell^j\end{bmatrix}^\top \in \mathbb{C}^\ell$, $j = 1 : \ell$. Then for $j = 1 : \ell$, we show that the spectral projection P_j is given by

$$P_j = \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j = \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI - A)^{-1}dz/\alpha_j,$$

where $q_{v_j}(z) := \begin{bmatrix}1 & z & \cdots & z^{\ell-1}\end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial determined by v_j and $\alpha_j := q_{v_j}(\lambda_j)$.

Several contour integration based methods have been proposed in the literature for solving linear and nonlinear matrix eigenvalue problems, see [51, 5, 11, 55, 54] and the references therein. However, none of these methods recovers algebraic multiplicities of the eigenvalues as well as the corresponding spectral projections. Moreover, the size of the eigenvalue problem required to be solved for computing $\lambda_1, \ldots, \lambda_\ell$ depend on the total algebraic multiplicity $m := m_1 + \cdots + m_\ell$ of the eigenvalues. By contrast, the size of the Hankel eigenvalue problem $(\widehat{H}_\ell - \lambda H_\ell)v = 0$ to be solved in our method for computing ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ is equal to ℓ irrespective of the algebraic multiplicities of the eigenvalues.

5.2 Moment problems for matrices

For simplicity of presenting the main ideas, we consider the special case of a matrix eigenvalue problem. Let $A \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathrm{eig}(A)$. Then the nonzero vectors v and u in $\mathbb{C}^n$ are said to be right and left eigenvectors of A corresponding λ if

$$Av = \lambda v \text{ and } u^\top A = \lambda u^\top.$$

Consider a matrix pencil $A - zB$, where A and B are matrices in $\mathbb{C}^{n\times n}$. The pencil $A - zB$ is said to be regular if $\det(A - \lambda B) \neq 0$ for some $\lambda \in \mathbb{C}$. Consider the generalized eigenvalue problem

$$(A - \lambda B)v = 0$$

where $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$ is nonzero. The vector v is called an eigenvector (or right eigenvector) of the pencil $A - zB$ corresponding to the eigenvalue λ. A nonzero vector $u \in \mathbb{C}^n$ is said to be a left eigenvector of $(A - zB)$ corresponding to λ if $u^\top(A - \lambda B) = 0$. If B is singular then ∞ is also an eigenvalue of $A - zB$. We denote the set of eigenvalues of the pencil $A - zB$ by $\mathrm{eig}(A, B)$.

5.2.1 Tracial moments and eigenvalues

Let $\Omega \subset \mathbb{C}$ be a bounded domain such that $\Gamma := \partial\Omega$ is a positively oriented rectifiable simple closed curve and that $\Gamma \subset \rho(A)$.

Theorem 5.2.1. *Suppose that* $\sigma := \mathrm{eig}(A) \cap \Omega = \{\lambda_1, \ldots, \lambda_\ell\}$. *Let* m_j *be the algebraic multiplicity of* λ_j *for* $j = 1 : \ell$. *Then for any analytic function* f *(analytic in an open set containing* σ*), we have*

$$s_f := \frac{1}{2\pi i} \int_\Gamma f(z)\, \mathrm{Tr}\left((zI - A)^{-1}\right) dz = m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell).$$

In particular, we have

$$\mathrm{mean}(\sigma) := \frac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell} = \frac{\int_\Gamma z\, \mathrm{Tr}\left((zI - A)^{-1}\right) dz}{\int_\Gamma \mathrm{Tr}\left((zI - A)^{-1}\right) dz}.$$

Proof. Let $\lambda_1, \ldots, \lambda_\ell, \ldots, \lambda_p$ be distinct eigenvalues of A with algebraic multiplicities $m_1, ..., m_\ell, \ldots, m_p$, respectively. Thus $\text{eig}(A) = \{\lambda_1, \ldots, \lambda_p\}$ and

$$\text{Tr}((zI - A)^{-1}) = \frac{m_1}{z - \lambda_1} + \cdots + \frac{m_\ell}{z - \lambda_\ell} + \cdots + \frac{m_p}{z - \lambda_p}.$$

Hence for any analytic function f (analytic in an open set containing σ), by Cauchy's integral formula [58], we have

$$\begin{aligned} s_f &:= \frac{1}{2\pi i} \int_\Gamma f(z)\, \text{Tr}((zI - A)^{-1}) dz \\ &= \frac{1}{2\pi i} \int_\Gamma f(z) \left(\frac{m_1}{z - \lambda_1} + \cdots \frac{m_\ell}{z - \lambda_\ell} + \cdots \frac{m_p}{z - \lambda_p} \right) dz \\ &= \frac{1}{2\pi i} \int_\Gamma f(z) \left(\frac{m_1}{z - \lambda_1} + \cdots + \frac{m_\ell}{z - \lambda_\ell} \right) dz \\ &= m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell). \end{aligned}$$

Now considering $f(z) := z$ and $f(z) := 1$, we have

$$\text{mean}(\sigma) := \frac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell} = \frac{\int_\Gamma z\, \text{Tr}((zI - A)^{-1}) dz}{\int_\Gamma \text{Tr}((zI - A)^{-1}) dz}.$$

□

We define the p-th order *tracial moment* of A associated with σ by

$$s_p := \frac{1}{2\pi i} \int_\Gamma z^p\, \text{Tr}((zI - A)^{-1}) dz, \quad p \in \mathbb{N} \cup \{0\}.$$

Then for $f_p(z) := z^p$, we have

$$s_p = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p, \quad p \in \mathbb{N} \cup \{0\}. \tag{5.1}$$

For $k \in \mathbb{N}$, we consider the $k \times k$ Hankel matrices H_k and $\widehat{H}_k$ given by

$$H_k := \text{Hankel}(s_0, \ldots, s_{2k-2}) \text{ and } \widehat{H}_k := \text{Hankel}(s_1, \ldots, s_{2k-1}).$$

Also consider the $k \times \ell$ Vandermonde matrix $V_k := \begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{k-1} & \ldots & \lambda_\ell^{k-1} \end{bmatrix}$ for $k \in \mathbb{N}$. Note that V_ℓ is invertible and $\text{rank}(V_k) = \ell$ for all $k \geq \ell$.

Theorem 5.2.2. *Suppose that* $\sigma := \mathrm{eig}(A) \cap \Omega = \{\lambda_1, \ldots, \lambda_\ell\}$*. Let* $m_1, \ldots, m_\ell$ *be the algebraic multiplicities of the eigenvalues* $\lambda_1, \ldots, \lambda_\ell$*, respectively. Then the Hankel matrix* H_ℓ *is nonsingular. Further,* $\mathrm{eig}(\widehat{H}_\ell, H_\ell) = \{\lambda_1, \ldots, \lambda_\ell\}$ *and* $\lambda_1, ..., \lambda_\ell$ *are simple eigenvalues of the* $\ell \times \ell$ *Hankel pencil* $\widehat{H}_\ell - \lambda H_\ell$*. The algebraic multiplicities* $m_1, \ldots, m_\ell$ *are given by unique solution of the Vandermonde system*

$$V_\ell \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}.$$

The number ℓ *satisfies* $\ell = \mathrm{rank}(H_\ell) = \mathrm{rank}(H_m) = \mathrm{rank}(H_{\ell+p})$ *for any non-negative integer* p*, where* $m := s_0 = m_1 + \cdots + m_\ell$*.*

Proof. Define $D_\ell := \mathrm{diag}(m_1, \ldots, m_\ell)$ and $\Lambda_\ell := \mathrm{diag}(\lambda_1, \ldots, \lambda_\ell)$. By (6.4), we have $s_p = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p$ for $p = 0, 1, \ldots, \ell-1$ which gives $V_\ell \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}$.

Again by (6.4) the (i, j)-th entry of H_k is given by (here $e_i, e_j \in \mathbb{C}^k$)

$$e_i^\top H_k e_j = s_{i+j-2} = m_1\lambda_1^{i+j-2} + \cdots + m_\ell\lambda_\ell^{i+j-2} \tag{5.2}$$

$$= \begin{bmatrix} \lambda_1^{i-1} & \cdots & \lambda_\ell^{i-1} \end{bmatrix} D_\ell \begin{bmatrix} \lambda_1^{j-1} \\ \vdots \\ \lambda_\ell^{j-1} \end{bmatrix} = e_i^\top V_k D_\ell V_k^\top e_j \tag{5.3}$$

which shows that $H_k = V_k D_\ell V_k^\top$ for $k \in \mathbb{N}$. Similarly, $\widehat{H}_k = V_k D_\ell \Lambda_\ell V_k^\top$ for $k \in \mathbb{N}$.

Thus we have $H_\ell = V_\ell D_\ell V_\ell^\top$ and $\widehat{H}_\ell = V_\ell D_\ell \Lambda_\ell V_\ell^\top$. Since V_ℓ and D_ℓ are invertible, the matrix H_ℓ is invertible. Now

$$\widehat{H}_\ell - \lambda H_\ell = V_\ell D_\ell (\Lambda_\ell - \lambda I_\ell) V_\ell^\top \tag{5.4}$$

shows that $\mathrm{eig}(\widehat{H}_\ell, H_\ell) = \{\lambda_1, \ldots, \lambda_\ell\}$ and that $\lambda_1, \ldots, \lambda_\ell$ are simple eigenvalues of the pencil $\widehat{H}_\ell - \lambda H_\ell$.

Finally, since $H_k = V_k D_\ell V_k^\top$ and $\mathrm{rank}(V_k) = \mathrm{rank}(V_\ell) = \ell$ for all $k \geq \ell$, it follows that $\mathrm{rank}(H_{\ell+p}) = \mathrm{rank}(H_m) = \mathrm{rank}(H_\ell) = \ell$ for all non-negative integer p. ☐

Remark 5.2.3. *Theorem 6.2.3 provides a method for computing the eigenvalues* $\lambda_1, \ldots, \lambda_\ell$ *of* A*. First, compute the tracial moment* s_0*, which gives the total number of eigenvalues (counting multiplicity) of* A *inside the curve* Γ*, that is, in* Ω*. Second, compute* $\ell := \mathrm{rank}(H_{s_0})$*, which gives the number of distinct eigenvalues of* A *inside the curve* Γ*. Finally, solve the* $\ell \times \ell$ *Hankel eigenvalue problem* $(\widehat{H}_\ell - \lambda H_\ell)v = 0$*.*

Remark 5.2.4. *Let* $H_\ell = U\Sigma V^*$ *be a singular value decomposition (SVD) of* H_ℓ*. Define* $A_\ell := \Sigma^{-1/2}U^*\widehat{H}_\ell V\Sigma^{-1/2}$*. Then the generalized eigenproblem* $(\widehat{H}_\ell - \lambda H_\ell)v = 0$ *reduces to the standard eigenproblem* $A_\ell u = \lambda u$*, where* $u := \Sigma^{1/2}V^*v$*. Indeed, we have*

$$(\widehat{H}_\ell - \lambda H_\ell)v = 0 \iff (\Sigma^{-1/2}U^*\widehat{H}_\ell V\Sigma^{-1/2} - \lambda I_\ell)\Sigma^{1/2}V^*v = 0.$$

Remark 5.2.5. *Observe that the Hankel pencil* $\widehat{H}_\ell - \lambda H_\ell$ *is complex symmetric. If the eigenvalues* $\lambda_1, \ldots, \lambda_\ell$ *are all real then by (6.4) the pencil* $\widehat{H}_\ell - \lambda H_\ell$ *is real symmetric. In particular, if the matrix* A *is Hermitian then the pencil* $\widehat{H}_\ell - \lambda H_\ell$ *is real symmetric. Suppose that the pencil* $\widehat{H}_\ell - \lambda H_\ell$ *is real symmetric. Let*

$$H_\ell = U\mathrm{diag}(\mu_1, \ldots, \mu_\ell)U^*$$

be a spectral decomposition of H_ℓ*. Set* $J_\ell := \mathrm{diag}(\mathrm{sign}(\mu_1), \ldots, \mathrm{sign}(\mu_\ell))$*. Then* J_ℓ *is a diagonal matrix with diagonal entries* ± 1*. Let* $D := \mathrm{diag}(|\mu_1|, \cdots, |\mu_\ell|)$*. Define*

$$A_\ell := J_\ell D^{-1/2}U^*\widehat{H}_\ell U D^{-1/2}.$$

Then A_ℓ *is* J_ℓ*-symmetric (i.e., pseudo-symmetric), that is,* $J_\ell A_\ell$ *is real symmetric. Consequently, the generalized eigenproblem* $(\widehat{H}_\ell - \lambda H_\ell)v = 0$ *reduces to the* J_ℓ*-symmetric standard eigenproblem* $A_\ell u = \lambda u$*, where* $u := D^{1/2}U^*v$*. Indeed,*

$$(\widehat{H}_\ell - \lambda H_\ell)v = 0 \iff (J_\ell D^{-1/2}U^*\widehat{H}_\ell U D^{-1/2} - \lambda I_\ell)D^{1/2}U^*v = 0.$$

Thus results available for pseudo-symmetric matrices can be utilized for the pseudo-symmetric matrix A_ℓ*.*

Approximation. Note that the tracial moment s_p for $p \in \mathbb{N} \cup \{0\}$ can be computed approximately by employing numerical integration. Given nodes $z_1, \ldots, z_N$ and weights $w_1, \ldots, w_N$, consider the numerical quadrature

$$s_p = \frac{1}{2\pi i}\int_\Gamma z^p \mathrm{Tr}((zI - A)^{-1})dz \approx \sum_{j=1}^{N} w_j z_j^p \mathrm{Tr}((z_j I - A)^{-1}). \tag{5.5}$$

Such a quadrature is obtained by considering a parametrization of Γ. Indeed, if $\phi : [a,b] \longrightarrow \mathbb{C}$ is a parametrization of the curve Γ then we have

$$s_p = \frac{1}{2\pi i}\int_a^b \phi(t)^p \mathrm{Tr}((\phi(t)I - A)^{-1})\phi'(t)dt,$$

where $\phi'(t)$ is the derivative of $\phi(t)$. Now, a numerical quadrature applied to the parametrized integral for s_p yields a numerical quadrature of the form (5.5) for s_p. In particular, if Γ is a circle with centre at c and radius r then $\phi(t) := c + re^{it}$, $t \in [0, 2\pi]$. Next, partition $[0, 2\pi]$ into N equal subintervals and consider the nodes $t_j := 2\pi j/N$ for $j = 0 : N$. This yields the nodes $z_j := \phi(t_j) = c + \exp(2\pi ij/N)$ for $j = 0 : N$ on the circle Γ. Since $z_0 = z_N$, the trapezoid rule with weights $w_0 = w_N = \pi/N$ and $w_j = 2\pi/N$, $j = 1 : N-1$, yields the trapezoid quadrature for the contour integral

$$\begin{aligned} s_p &= \frac{1}{2\pi i}\int_\Gamma z^p \mathrm{Tr}((zI - A)^{-1})dz = \frac{1}{2\pi}\int_0^{2\pi} \phi(t)^p \mathrm{Tr}((\phi(t)I - A)^{-1})re^{it}dt \\ &\approx \frac{1}{2\pi}\sum_{j=0}^{N} w_j z_j^p \mathrm{Tr}((z_jI - A)^{-1})\, r\exp(2\pi ij/N) \\ &= \frac{1}{N}\sum_{j=1}^{N} z_j^p \mathrm{Tr}((z_jI - A)^{-1})\, r\exp(2\pi ij/N). \end{aligned}$$

The approximate tracial moments, that is, approximation of s_p obtained by a numerical quadrature yields the following algorithm.

Input: Nodes z_j and weights w_j for j$= 1 : N$ as in (5.5) and a matrix A.
Output: Eigenpair (λ_j, v_j) of $\hat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$.

Compute $\mathtt{s_0} = \mathtt{sum}_{\mathtt{j=0}}^{\mathtt{N}} \mathtt{w_j}\ \mathrm{Tr}((\mathtt{z_j I} - \mathtt{A})^{-1})$

Compute $\mathtt{s_p} := \mathtt{sum}_{\mathtt{j=0}}^{\mathtt{N}} \mathtt{w_j}\ \mathtt{z_j^p}\ \mathrm{Tr}((\mathtt{z_j I} - \mathtt{A})^{-1})$, $\mathtt{p} = 1{:}\,2\mathtt{s_0} - 2$

Compute the Hankel matrix $\mathtt{H_{s_0}}$ and compute $\ell = \mathrm{rank}(\mathtt{H_{s_0}})$

Compute the $\ell \times \ell$ Hankel matrices $\mathtt{H}_\ell$ and $\widehat{\mathtt{H}}_\ell$

Compute eigenvalues and eigenvectors of the pencil $\widehat{\mathtt{H}}_\ell - \lambda\mathtt{H}_\ell$

We now illustrate performance of the above algorithm by considering an example.

Example 5.2.6. *Consider the* 5×5 *matrix* `A = gallery(5)` *from* MATLAB *gallery. It* *is known that A is a nilpotent matrix and is given by*

$$A = \begin{bmatrix} -9 & 11 & -21 & 63 & -252 \\ 70 & -69 & 141 & -421 & 1684 \\ -575 & 575 & -1149 & 3451 & -13801 \\ 3891 & -3891 & 7782 & -23345 & 93365 \\ 1024 & -1024 & 2048 & -6144 & 24572 \end{bmatrix}.$$

The matrix A has only one eigenvalue $\lambda := 0$ with geometric multiplicity 1 and algebraic multiplicity 5. The MATLAB *command* `eig(A)` *computes the following eigenvalues*

```
-3.472940132398842e-02 + 2.579009841174434e-02i
-3.472940132398842e-02 - 2.579009841174434e-02i
 1.377760760018629e-02 + 4.011025813393478e-02i
 1.377760760018629e-02 - 4.011025813393478e-02i
 4.190358744843689e-02 + 0.000000000000000e+00i
```

Observe that the computed eigenvalues almost lie on a circle of radius $\mathcal{O}(10^{-2})$ centred at 0. If $\lambda \in \mathbb{C}$ then the smallest singular value $\sigma_{\min}(A - \lambda I)$ of $A - \lambda I$ is called the backward error of λ as an approximate eigenvalue of A. The errors and backward errors of the computed eigenvalues are given in the following table.

```
     Error          Backward error
-------------------------------
4.3258e-02         5.7812e-13
4.3258e-02         5.7812e-13
4.2411e-02         5.3920e-13
4.2411e-02         5.3920e-13
4.1904e-02         5.6750e-13
```

MATLAB *ensures that the computed eigenvalues are exact eigenvalues of $A + E$ for some matrix E such that $\|A\|_2/\|E\|_2 = \mathcal{O}(10^{-16})$. Since A is similar to 5×5 nilpotent Jordan block, the eigenvalue $\lambda := 0$ splits into five simple eigenvalues $\lambda_1, \ldots, \lambda_5$ of $A+E$. It is well known [33] that $|\lambda - \lambda_j| = \mathcal{O}(\|E\|_2^{1/5})$. This is reflected in the errors given in the first column of the above table.*

We now compute the eigenvalue of A using the above mentioned algorithm. We consider Γ to be the unit circle $\phi(t) := e^{it}$, $t \in [0, 2\pi]$. Then for the nodes $z_j :=$

$\exp(2\pi ij/N)$, $j = 0 : N$, *on the unit circle* Γ, *the trapezoid quadrature is given by*

$$\begin{aligned} s_p &= \frac{1}{2\pi i}\int_\Gamma z^p \mathrm{Tr}((zI - A)^{-1})dz \approx \frac{1}{N}\sum_{j=1}^{N} z_j^p \mathrm{Tr}((z_jI - A)^{-1})\exp(2\pi ij/N) \\ &= \frac{1}{N}\sum_{j=1}^{N} z_j^{p+1}\mathrm{Tr}((z_jI - A)^{-1}),\ p \in \mathbb{N} \cup \{0\}. \end{aligned}$$

For $N := 100$, *we have the following results. Although,* $\ell = \mathrm{rank}(H_5) = 1$, *the approximate Hankel matrix* $\tilde{H}_5$ *obtained from approximate tracial moments has numerical rank* 5. *Indeed, the singular values of* $\tilde{H}_5$ *are given by*

```
5.0000e+00, 1.1653e-06, 4.7683e-07, 1.6688e-07, 8.9832e-08
```

which show that the (numerical) rank is 5. *The last four singular values are relatively small and are results of approximation errors in* $\tilde{H}_5$. *The eigenvalues of the* 5×5 *approximate Hankel pencil and their multiplicities are given in the following table.*

Eigenvalues of Hankel pencil	Multiplicity
5.247683437289830e-01 + 5.582969225138145e-01i	0
5.248101752705502e-01 - 5.582894562824255e-01i	0
-5.409096854206018e-01 + 3.395871670942288e-01i	0
-5.409116221622252e-01 - 3.395993770642578e-01i	0
2.347822865490909e-07 - 1.825809837463345e-11i	5

The multiplicity of first four computed eigenvalues is 0 *which indicates that these are spurious eigenvalues. Thus the algorithm correctly detects the spurious eigenvalues, as it should, since the Hankel pencil* $\widehat{H}_1 - \lambda H_1 = s_1 - \lambda s_0$ *has only one eigenvalue* $\lambda = s_1/s_0 = 0$. *The errors and backward errors of the computed eigenvalues are given in the following table.*

```
 Error          Backward error
-----------------------------
7.6621e-01      2.5809e-06
7.6623e-01      2.5816e-06
6.3867e-01      1.4491e-07
6.3868e-01      1.4492e-07
2.3478e-07      5.2638e-15
```

Note that errors as well as backward errors of the first four spurious eigenvalues are relatively large. Discarding the spurious eigenvalues, we obtain only one computed eigenvalue. Observe that the error in the computed eigenvalue is $\mathcal{O}(10^{-7})$ *and the backward error is* $\mathcal{O}(10^{-15})$*, which are far better than those of the computed eigenvalues obtained by* MATLAB *command* `eig(A)`.

5.2.2 Matrix moments and spectral projection

We now describe the recovery of spectral projections (and hence eigenvectors) of A from those of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. Recall that A has ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ inside the curve Γ and $\sigma := \{\lambda_1, \ldots, \lambda_\ell\}$. Note that

$$P := \frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} dz$$

is the spectral projection of A corresponding to the eigenvalues in σ (see, Section 5.4). Let P_j denote the spectral projection of A corresponding to λ_j for $j = 1 : \ell$. Then

$$P = P_1 + \cdots + P_\ell \quad \text{and} \quad P_i P_j = 0 \quad \text{for} \quad i \neq j.$$

Further, we have $R(P_j) = N((A - \lambda_j I)^{m_j})$. Note that $P^\top$ is the spectral projection of $A^\top$ corresponding to σ and $R(P_j^\top) = N((A^\top - \lambda_j I)^{m_j})$. In the special case, when λ_j is semisimple, we have $R(P_j) = N(A - \lambda_j I)$ and $R(P_j^\top) = N((A - \lambda_j I)^\top)$.

Now consider the *matrix moments*

$$S_p := \frac{1}{2\pi i}\int_\Gamma z^p (zI - A)^{-1} dz, \quad p = 0, 1, \ldots, \ell - 1. \tag{5.6}$$

Let $X \in \mathbb{C}^{n\times n}$ be such that $X^{-1}AX = \operatorname{diag}(A_1, \ldots, A_\ell, \widehat{A})$, where $\operatorname{eig}(A_j) = \{\lambda_j\}$ for $j = 1 : \ell$ and $\operatorname{eig}(\widehat{A}) = \operatorname{eig}(A) \setminus \sigma$. Let

$$X = \begin{bmatrix} X_1 & \cdots & X_\ell & \widehat{X} \end{bmatrix} \text{ and } Y := (X^{-1})^\top = \begin{bmatrix} Y_1 & \cdots & Y_\ell & \widehat{Y} \end{bmatrix}$$

be conformal partitions of X and Y. Then we have

$$(A - zI)^{-1} = \sum_{j=1}^{\ell} X_j(A_j - zI_{m_j})^{-1}Y_j^\top + \widehat{X}(\widehat{A} - zI)^{-1}\widehat{Y}^\top. \tag{5.7}$$

Hence by (5.6) and (5.7), we have

$$S_p = \sum_{j=1}^{\ell} X_j A_j^p Y_j^\top, \quad P = \sum_{j=1}^{\ell} X_j Y_j^\top \text{ and } P_j = X_j Y_j^\top \text{ for } j = 1 : \ell. \tag{5.8}$$

By Theorem 6.2.3, all eigenvalues of $\widehat{H}_\ell - \lambda H_\ell$ are simple and $\text{eig}(\widehat{H}_\ell, H_\ell) = \sigma$. Thus, an eigenvector v_j of $\widehat{H}_\ell - \lambda H_\ell$ corresponding to λ_j, that is, $(\widehat{H}_\ell - \lambda_j H_\ell)v_j = 0$, is unique up to a scalar multiple. Note that (λ_j, v_j) is also a left eigenpair of $\widehat{H}_\ell - \lambda H_\ell$.

Observe that $S_0 = P$ is the spectral projection and $R(P)$ is the spectral subspace of A corresponding to the eigenvalues in σ. Now we extract the individual spectral projection P_j from the ℓ matrix moments $S_0, \ldots, S_{\ell-1}$, which can be utilized to recover eigenvectors of A from those of $\widehat{H}_\ell - \lambda H_\ell$.

Let $v \in \mathbb{C}^\ell$ be given by $v := \begin{bmatrix} v_1 & \cdots & v_\ell \end{bmatrix}^\top$. Then v defines a polynomial $q_v(z) \in \mathbb{C}[z]$ of degree at most $\ell - 1$ given by

$$q_v(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v = v_1 + v_2 z + \cdots + v_\ell z^{\ell-1}.$$

Theorem 5.2.7. *Let (λ_j, v_j) be an eigenpair of $\widehat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$. Define*

$$\mathbf{V} := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \in \mathbb{C}^{n \times n\ell} \text{ and } \mathbf{U} := \begin{bmatrix} S_0^\top & \cdots & S_{\ell-1}^\top \end{bmatrix} \in \mathbb{C}^{n \times n\ell}.$$

Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A are semisimple. Set $\alpha_j := q_{v_j}(\lambda_j)$ for $j = 1 : \ell$. Then for $j = 1 : \ell$,

$$P_j = \mathbf{V}(v_j \otimes I_n)/\alpha_j \text{ and } P_j^\top = \mathbf{U}(v_j \otimes I_n)/\alpha_j$$

are spectral projections of A and $A^\top$ corresponding to λ_j, respectively.

Define $\mathbf{v}_j := \mathbf{V}(v_j \otimes v) = \begin{bmatrix} S_0 v \cdots S_{\ell-1} v \end{bmatrix} v_j$ and $\mathbf{u}_j := \mathbf{U}(v_j \otimes u) = \begin{bmatrix} S_0^\top u \cdots S_{\ell-1}^\top u \end{bmatrix} v_j$ for $v, u \in \mathbb{C}^n$. Then $\mathbf{v}_j$ and $\mathbf{u}_j$, respectively, are right and left eigenvectors of A corresponding to λ_j whenever $\mathbf{v}_j \neq 0$ and $\mathbf{u}_j \neq 0$.

Proof. Since λ_j is semisimple, the matrix A_j in (5.7) can be chosen as $A_j = \lambda_j I_{m_j}$. Consequently, by (5.8) we have

$$S_p = \sum_{j=1}^{\ell} \lambda_j^p X_j Y_j^\top = \sum_{j=1}^{\ell} \lambda_j^p P_j = \begin{bmatrix} P_1 & \cdots & P_\ell \end{bmatrix} \begin{bmatrix} \lambda_1^p \otimes I_n \\ \vdots \\ \lambda_\ell^p \otimes I_n \end{bmatrix}.$$

This shows that $\mathbf{V} = \begin{bmatrix} P_1 & \cdots & P_\ell \end{bmatrix} (V_\ell^\top \otimes I_n) \Longrightarrow P_j = \mathbf{V}(V_\ell^{-\top} e_j \otimes I_n)$, where V_ℓ is the $\ell \times \ell$ Vandermonde matrix given in Theorem 6.2.3. Now by (6.5), we have $(\Lambda_\ell - \lambda_j I_\ell) V_\ell^\top v_j = 0$ which shows that $V_\ell^\top v_j = \alpha_j e_j \Longrightarrow V_\ell^{-\top} e_j = v_j/\alpha_j$ for $j = 1 : \ell$. Hence we have $P_j = \mathbf{V}(v_j \otimes I_n)/\alpha_j$ for $j = 1 : \ell$. The proof for $P_j^\top$ is immediate.

Since $AP_j = \lambda_j P_j$, it follows that $P_j v$ is an eigenvector of A corresponding to λ_j provided that $P_j v \neq 0$. Hence $\alpha_j P_j v = \mathbf{V}(v_j \otimes v) = \begin{bmatrix} S_0 v \cdots S_{\ell-1} v \end{bmatrix} v_j = \mathbf{v}_j$ is a right eigenvector of A corresponding to λ_j whenever $\mathbf{v}_j \neq 0$. Similarly, it follows that $\alpha_j P_j^\top u = \mathbf{U}(v_j \otimes u) = \begin{bmatrix} S_0^\top u \cdots S_{\ell-1}^\top u \end{bmatrix} v_j = \mathbf{u}_j$ is a left eigenvector of A corresponding to λ_j whenever $\alpha_j P_j^\top v = \mathbf{u}_j \neq 0$. □

Corollary 5.2.8. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of A. Let P_j be the spectral projection of A corresponding to λ_j and let v_j be given by $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then for $j = 1 : \ell$,*

$$\begin{aligned} P_j &= \mathbf{V}(v_j \otimes I_n)/\alpha_j = \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI_n - A)^{-1} dz\right)/\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial and $\alpha_j := q_{v_j}(\lambda_j)$.

Proof. By (5.6), we have

$$\begin{aligned} \mathbf{V}(v_j \otimes I_n) &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right) \\ &= \frac{1}{2\pi i}\int_\Gamma \sum_{k=1}^{\ell} \beta_k^j z^{k-1}(zI_n - A)^{-1} dz \\ &= \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI_n - A)^{-1} dz, \end{aligned}$$

which yields the desired result. □

Theorem 5.2.7 shows that exactly ℓ matrix moments $S_0, \ldots, S_{\ell-1}$ are needed for recovering the ℓ spectral projections $P_1, \ldots, P_\ell$ of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively. Also, for a random choice of u and v in $\mathbb{C}^n$, the vectors $\mathbf{V}(v_j \otimes v)$ and $\mathbf{U}(v_j \otimes u)$ are likely to be nonzero. Hence each such pair of vectors would yield a pair of right and left eigenvectors of A corresponding to λ_j. On the other hand, if all linearly independent right and left eigenvectors of A corresponding to λ_j are required then rank revealing QR factorization of $\mathbf{V}(v_j \otimes v)$ and $\mathbf{U}(v_j \otimes u)$ can be gainfully utilized.

Remark 5.2.9. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be simple eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalue of A. Then by Corollary 5.2.8, we have*

$$\mathbf{V}(v_j \otimes I_n) = \beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1} \in \mathbb{C}^{n\times n}$$

and $\mathrm{rank}(\mathbf{V}(v_j \otimes I_n)) = \mathrm{rank}(P_j) = m_j$ *for* $j = 1 : \ell$. *Let* $\mathbf{V}(v_j \otimes I_n)\widehat{P} = Q \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}$ *be a rank revealing QR factorization, where $\widehat{P} \in \mathbb{C}^{n\times n}$ is a permutation matrix, $Q \in \mathbb{C}^{n\times n}$ is unitary and $R_{11} \in \mathbb{C}^{m_j \times m_j}$ is a nonsingular upper triangular matrix. Then the first m_j columns of Q form an orthonormal basis of $N(A - \lambda_j I_n)$.*

A better method to recover all linearly independent right and left eigenvectors of A from eigenvectors of the pencil $\widehat{H}_\ell - \lambda H_\ell$ is to consider a compact SVD of $\mathbf{V}(v_j \otimes I_n)$ for $j = 1 : \ell$. We have the following result.

Theorem 5.2.10. *Let (λ_j, v_j) be an eigenpair of $\widehat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$. Define $\mathbf{V} := \begin{bmatrix} S_0 & \cdots & S_\ell \end{bmatrix} \in \mathbb{C}^{n\times n\ell}$. Consider a compact SVD of $\mathbf{V}(v_j \otimes I_n)$ given by*

$$\mathbf{V}(v_j \otimes I_n) = X\Sigma Y^* \quad \text{where } X \in \mathbb{C}^{n\times m_j} \text{ and } Y \in \mathbb{C}^{n\times m_j} \text{ are isometry.}$$

Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of A. Then the columns of X form an orthonormal basis of $N(A - \lambda_j I_n)$ and the columns of Y form an orthonormal basis of $N((A - \lambda_j I_n)^*)$.

Proof. By Theorem 5.2.7, we have $N(A - \lambda_j I_n) = R(P_j) = R(\mathbf{V}(v_j \otimes I_n))$. Note that $\mathrm{rank}(P_j) = m_j$ for $j = 1 : \ell$. Let $\mathbf{V}(v_j \otimes I_n) = X\Sigma Y^*$ be a compact SVD of $\mathbf{V}(v_j \otimes I_n)$, where $X \in \mathbb{C}^{n\times m_j}$ and $Y \in \mathbb{C}^{n\times m_j}$ are isometry and $\Sigma \in \mathbb{C}^{m_j\times m_j}$ is a diagonal matrix containing m_j nonzero singular values of $\mathbf{V}(v_j \otimes I_n)$. Then we have

$$\begin{aligned} R(X) &= R(\mathbf{V}(v_j \otimes I_n)) = R(P_j) = N(A - \lambda_j I_n) \text{ and} \\ R(Y) &= R((\mathbf{V}(v_j \otimes I_n))^*) = R(P_j^*) = N((A - \lambda_j I_n)^*). \end{aligned}$$

□

Observe that the eigenvectors $v_1, \ldots, v_\ell$ have not been assumed to be normalized. For computation, it may be advisable to normalize the eigenvectors. For instance, we could normalize the eigenvectors either to have unit norm or satisfy $q_{v_j}(\lambda_j) = 1$ for $j = 1 : \ell$. Also, notice that a compact SVD of $\mathbf{V}(v_j \otimes I_n)$ yields both right and left eigenvectors of A corresponding to λ_j. In fact, it yields orthonormal bases of $N(A - \lambda_j I_n)$

and $N(A-\lambda_j I_n)^*)$. Thus, in view of Corollary 5.2.8, the recovery method for P_j yields a practical algorithm when $\mathbf{V}(v_j \otimes I_n)$ is approximated by a numerical quadrature:

$$\begin{aligned}\mathbf{V}(v_j \otimes I_n) &= \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI_n - A)^{-1}dz \\ &\approx \sum_{k=1}^{N} w_k q_{v_j}(z_k)(z_k I_n - A)^{-1}\end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1}\end{bmatrix} v_j \in \mathbb{C}[z]$ for $j = 1 : \ell - 1$.

Input: Nodes z_j and weights w_j for j= $1 : N$ as in (5.5) and a matrix A.
Eigenpair (λ_j, v_j) of $\hat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$.
Output: Orthonormal bases of $N(A-\lambda_j I_n)$ and $N((A-\lambda_j I_n)^*)$ for $j = 1 : \ell$.

For j=1:ℓ

Compute $\mathtt{q_{v_j}(z_k)} = \begin{bmatrix} \mathtt{1} & \mathtt{z_k} & \cdots & \mathtt{z_k^{\ell-1}}\end{bmatrix} \mathtt{v_j}$ `for k = 1:N`

Compute $\mathtt{A_j} = \mathtt{sum}_{\mathtt{k=1}}^{\mathtt{N}} \mathtt{w_k}\, \mathtt{q_{v_j}(z_k)(z_k I - A)^{-1}}$

Compute compact SVD $[\mathtt{X_j}, \mathtt{D_j}, \mathtt{Y_j}] = \mathtt{svd}(\mathtt{A_j})$

Return the isometries $\mathtt{X_j}$ and $\mathtt{Y_j}$

End

The algorithm returns the isometries X_j and Y_j such that $R(X_j) = N(A - \lambda_j I_n)$ and $R(Y_j) = N((A-\lambda_j I_n)^*)$ for $j = 1 : \ell$. Thus columns of X_j and Y_j, respectively, are orthonormal right and left eigenvectors of A corresponding to λ_j for $j = 1 : \ell$.

5.3 Moment problems for matrix pencils

The moment problems for matrix pencils can be solved in a similar manner by appropriately modifying the moments. We briefly discuss the main results. Let $A - \lambda B$ be an $n \times n$ regular pencil. Let $\lambda_1, \ldots, \lambda_p$ be distinct finite eigenvalues of the pencil $A - \lambda B$ with algebraic multiplicities $m_1, \ldots, m_p$, respectively. For $\ell \leq p$, assume that the eigenvalues $\lambda_1, ..., \lambda_\ell$ lie inside the region enclosed by the curve Γ, that is, $\sigma := \mathrm{eig}(A, B) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, ..., \lambda_\ell\}$. For $p \in \mathbb{N} \cup \{0\}$, we consider the moments

$$s_p := \frac{-1}{2\pi i}\int_\Gamma z^p\, \mathrm{Tr}((A - zB)^{-1}B)dz \text{ and } S_p := \frac{-1}{2\pi i}\int_\Gamma z^p\, (A - zB)^{-1}Bdz.$$

Theorem 5.3.1. *Let* $\sigma := \mathrm{eig}(A,B) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, ..., \lambda_\ell\}$. *Then for any analytic function* f *(analytic in an open set containing* σ*), we have*

$$s_f := \frac{-1}{2\pi i}\int_\Gamma f(z)\, \mathrm{Tr}((A - zB)^{-1}B)dz = m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell).$$

In particular, we have we have

$$\mathrm{mean}(\sigma) := \frac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell} = \frac{\int_\Gamma z\, \mathrm{Tr}((A - zB)^{-1}B)dz}{\int_\Gamma \mathrm{Tr}((A - zB)^{-1}B)dz}.$$

Proof. By Weierstrass canonical form, there exist $n \times n$ nonsingular matrices X and Y such that

$$Y^{-1}(A - \lambda B)X = \left[\begin{array}{c|c} J - \lambda I & 0 \\ \hline 0 & I - \lambda N \end{array}\right],$$

where J is in Jordan canonical form, $\mathrm{eig}(J) := \{\lambda_1, \ldots, \lambda_p\}$ and N is a nilpotent matrix in Jordan canonical form. This shows that $Y^{-1}BX = \mathrm{diag}(I, N)$. Hence

$$\begin{aligned} \mathrm{Tr}((A - \lambda B)^{-1}B) &= \mathrm{Tr}((J - \lambda I)^{-1}) + \mathrm{Tr}((I - \lambda N)^{-1}N) = \mathrm{Tr}((J - \lambda I)^{-1}) \\ &= \frac{m_1}{\lambda_1 - \lambda} + \cdots \frac{m_\ell}{\lambda_\ell - \lambda} + \cdots \frac{m_p}{\lambda_p - \lambda}. \end{aligned}$$

Since f is analytic, by Cauchy's integral formula, we have

$$\begin{aligned} s_f &:= \frac{-1}{2\pi i}\int_\Gamma f(z)\, \mathrm{Tr}((A - zB)^{-1})dz = \frac{1}{2\pi i}\int_\Gamma f(z)\, \mathrm{Tr}((zI - J)^{-1})dz \\ &= \frac{1}{2\pi i}\int_\Gamma f(z)\left(\frac{m_1}{z - \lambda_1} + \cdots \frac{m_\ell}{z - \lambda_\ell} + \cdots \frac{m_p}{z - \lambda_p}\right)dz \\ &= \frac{1}{2\pi i}\int_\Gamma f(z)\left(\frac{m_1}{z - \lambda_1} + \cdots + \frac{m_\ell}{z - \lambda_\ell}\right)dz \\ &= m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell). \end{aligned}$$

Next, taking $f(z) := z$ and $f(z) := 1$, we have

$$\mathrm{mean}(\sigma) := \frac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell} = \frac{\int_\Gamma z\, \mathrm{Tr}((A - zB)^{-1}B)dz}{\int_\Gamma \mathrm{Tr}((A - zB)^{-1}B)dz}.$$

This completes the proof. □

For $f_k(z) := z^k$, by Theorem 5.3.1, we have

$$s_k = \frac{-1}{2\pi i}\int_\Gamma z^k\, \mathrm{Tr}((A - zB)^{-1}B)dz = m_1\lambda_1^k + \cdots + m_\ell\lambda_\ell^k, \quad k = 0, 1, \ldots. \tag{5.9}$$

As before, construct the $k \times k$ Hankel matrix H_k and shifted Hankel matrix $\widehat{H}_k$ from the moments s_k for $k \in \mathbb{N}$. Then we have the following result which is an analogue of Theorem 6.2.3.

Theorem 5.3.2. *Suppose that $A - \lambda B$ has ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ inside the curve Γ and the multiplicities of the eigenvalues are $m_1, \ldots, m_\ell$, respectively. Then $\lambda_1, \ldots, \lambda_\ell$ are the simple eigenvalues of the $\ell \times \ell$ Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. The algebraic multiplicities $m_1, \ldots, m_\ell$ are given by the $\ell \times \ell$ Vandermonde system*

$$\begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}. \tag{5.10}$$

The number ℓ satisfies $\ell = \mathrm{rank}(H_\ell) = \mathrm{rank}(H_m) = \mathrm{rank}(H_{\ell+p})$ for any non-negative integer p, where $m := s_0 = m_1 + \cdots + m_\ell$.

Proof. A verbatim proof of Theorem 6.2.3 yields the desired results. □

Theorem 5.3.2 provides a method for computing the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of the pencil $A - \lambda B$. If the eigenvalues are real then by (5.9) the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ is real symmetric. In particular, if $A - \lambda B$ is a definite pencil, that is, $A^* = A$ and $B^* = B$ and $\min_{\|x\|_2=1}[(x^*Ax)^2 + x^*Bx)^2] > 0$, then the pencil $\widehat{H}_\ell - \lambda H_\ell$ is real symmetric.

Spectral projections. We now describe recovery of spectral projections (and hence eigenvectors) of $A - \lambda B$ from those of the pencil $\widehat{H}_\ell - \lambda H_\ell$. Define

$$P := \frac{-1}{2\pi i} \int_\Gamma (A - zB)^{-1} B dz \quad \text{and} \quad Q := \frac{-1}{2\pi i} \int_\Gamma B(A - zB)^{-1} dz. \tag{5.11}$$

Then P and Q, respectively, are called right and left spectral projections of $A - \lambda B$ corresponding to σ, see [23]. Further, $R(P)$ and $R(Q^\top)$, respectively, are right and left deflating subspaces of $A - \lambda B$. Let P_j and Q_j, respectively, be right and left spectral projections of $A - \lambda B$ corresponding to λ_j for $j = 1 : \ell$. Then

$$\begin{aligned} P &= P_1 + \cdots + P_\ell, \quad P_i P_j = 0 \ \text{ for } i \neq j, \\ Q &= Q_1 + \cdots + Q_\ell, \quad Q_i Q_j = 0 \ \text{ for } \ i \neq j. \end{aligned}$$

Let X and Z be nonsingular matrices such that

$$Z^{-1}(A - \lambda B)X = \mathrm{diag}(A_1 - \lambda I_{m_1}, \ldots, A_\ell - \lambda I_{m_\ell}, \widehat{A} - \lambda \widehat{B}),$$

where $\mathrm{eig}(A_j) = \{\lambda_j\}$, $j = 1 : \ell$, and $\mathrm{eig}(\widehat{A}, \widehat{B}) = \mathrm{eig}(A, B) \setminus \sigma$. Let

$$X = \begin{bmatrix} X_1 & \cdots & X_\ell & \widehat{X} \end{bmatrix} \ \text{ and } \ Y := (Z^{-1})^\top = \begin{bmatrix} Y_1 & \cdots & Y_\ell & \widehat{Y} \end{bmatrix}$$

be conformal partitions of X and Y. Then $Y_j^\top BX_j = I_{m_j}$, $j = 1 : \ell$, and $Y_i^\top BX_j = 0$ for $i \neq j$. Now, we have

$$(A - zB)^{-1} = \sum_{j=1}^{\ell} X_j(A_j - zI_{m_j})^{-1}Y_j^\top + \widehat{X}(\widehat{A} - z\widehat{B})^{-1}\widehat{Y}^\top. \qquad (5.12)$$

Hence by (5.11) and (5.12), we have

$$\begin{aligned} P &= X_1Y_1^\top B + \cdots + X_\ell Y_\ell^\top B, \quad P_j = X_jY_j^\top B \text{ for } j = 1 : \ell, \\ Q &= BX_1Y_1^\top + \cdots + BX_\ell Y_\ell^\top, \quad Q_j = BX_jY_j^\top \text{ for } j = 1 : \ell. \end{aligned}$$

If $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $A - \lambda B$ then we can choose X and Z such that $A_j = \lambda_j I_{m_j}, j = 1 : \ell$. Hence, in such a case, we have

$$R(X_j) = R(P_j) = N(A - \lambda_j B) \quad \text{and} \quad R(Y_j) = R(Q^\top) = N((A - \lambda_j B)^\top). \qquad (5.13)$$

Thus columns of X_j and Y_j are right and left eigenvectors of $A - \lambda B$ corresponding to λ_j for $j = 1 : \ell$.

Next, for $p = 0, 1, \ldots, \ell - 1$, consider the *matrix moments*

$$S_p := \frac{-1}{2\pi i}\int_\Gamma z^p(A - zB)^{-1}Bdz, \qquad M_p := \frac{-1}{2\pi i}\int_\Gamma z^pB(A - zB)^{-1}dz. \qquad (5.14)$$

Then by (5.12), we have $S_p = \sum_{j=1}^{\ell} X_jA_j^pY_j^\top B$ and $M_p = \sum_{j=1}^{\ell} BX_jA_j^pY_j^\top$. In particular, when $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $A - \lambda B$, we have

$$S_p = \sum_{j=1}^{\ell} \lambda_j^p X_j^p Y_j^\top B = \sum_{j=1}^{\ell} \lambda_j^p P_j \quad \text{and} \quad M_p = \sum_{j=1}^{\ell} \lambda_j^p BX_j^pY_j^\top = \sum_{j=1}^{\ell} \lambda_j^p Q_j. \qquad (5.15)$$

We have the following result which extracts the individual spectral projection P_j and Q_j from the matrix moments S_j and M_j for $j = 1, \ldots, \ell - 1$ and recovers eigenvectors of $A - \lambda B$ from those of $\widehat{H}_\ell - \lambda H_\ell$.

Theorem 5.3.3. *Let (λ_j, v_j) be an eigenpair of $\widehat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$. Define*

$$V := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \in \mathbb{C}^{n\times n\ell} \text{ and } U := \begin{bmatrix} M_0^\top & \cdots & M_{\ell-1}^\top \end{bmatrix} \in \mathbb{C}^{n\times n\ell}.$$

Set $\alpha_j := [1, \lambda_j, \ldots, \lambda_j^{\ell-1}]v_j$ for $j = 1 : \ell$. Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $A - \lambda B$ are semisimple. Then for $j = 1 : \ell$,

$$P_j = V(v_j \otimes I_n)/\alpha_j \text{ and } Q_j^\top = U(v_j \otimes I_n)/\alpha_j.$$

Define $\mathbf{v}_j := V(v_j \otimes v) = \begin{bmatrix} S_0v \cdots S_{\ell-1}v \end{bmatrix} v_j$ and $\mathbf{u}_j := U(v_j \otimes u) = \begin{bmatrix} M_0^\top u \cdots M_{\ell-1}^\top u \end{bmatrix} v_j$ for $v, u \in \mathbb{C}^n$. Then $\mathbf{v}_j$ and $\mathbf{u}_j$, respectively, are right and left eigenvectors of $A - \lambda B$ corresponding to λ_j whenever $\mathbf{v}_j \neq 0$ and $\mathbf{u}_j \neq 0$.

Proof. The proof for $P_j = V(v_j \otimes I_n)/\alpha_j$ is exactly the same as given in the proof of Theorem 5.2.7. Note that $\mathbf{v}_j = V(v_j \otimes v) = \alpha_j P_j v$. Hence it follows from (5.13) that $\mathbf{v}_j$ is a right eigenvector of $A - \lambda B$ corresponding to λ_j when $\alpha_j P_j v = \mathbf{v}_j \neq 0$.

Next, by (5.15) we have

$$M_p^\top = \sum_{j=1}^{\ell} \lambda_j^p (BX_jY_j^\top)^\top = \sum_{j=1}^{\ell} \lambda_j^p Q_j^\top = \begin{bmatrix} Q_1^\top & \cdots & Q_\ell^\top \end{bmatrix} \begin{bmatrix} \lambda_1^p \otimes I_n \\ \vdots \\ \lambda_\ell^p \otimes I_n \end{bmatrix}.$$

This shows that $U = \begin{bmatrix} Q_1^\top & \cdots & Q_\ell^\top \end{bmatrix} (V_\ell^\top \otimes I_n) \implies Q_j^\top = U(V_\ell^{-\top} e_j \otimes I_n)$, where V_ℓ is the $\ell \times \ell$ Vandermonde matrix given in Theorem 6.2.3. Now by (6.5), we have $(\Lambda_\ell - \lambda_j I_\ell)V_\ell^\top v_j = 0$ which shows that $V_\ell^\top v_j = \alpha_j e_j \implies V_\ell^{-\top} e_j = v_j/\alpha_j$ for $j = 1 : \ell$. Hence we have $Q_j^\top = U(v_j \otimes I_n)/\alpha_j$ for $j = 1 : \ell$. By (5.13) $\mathbf{u}_j = \alpha_j Q_j^\top u$ is a left eigenvector of $A - \lambda B$ corresponding to λ_j when $\alpha_j Q_j^\top u = \mathbf{u}_j \neq 0$. □

Given an eigenpair (λ_j, v_j) of $\widehat{H}_\ell - \lambda H_\ell$, for a random choice of v and u in $\mathbb{C}^n$, the vectors $V(v_j \otimes v)$ and $U(v_j \otimes u)$ are expected to be nonzero and each such choice would provide a pair of right and left eigenvectors of $A - \lambda B$ corresponding to λ_j.

A reliable option is to utilize compact SVD of $V(v_j \otimes I_n)$ and $U(v_j \otimes I_n)$. We have the following result.

Theorem 5.3.4. *Let (λ_j, v_j) be an eigenpair of $\widehat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$. Define*

$$V := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \in \mathbb{C}^{n \times n\ell} \text{ and } U := \begin{bmatrix} M_0^\top & \cdots & M_{\ell-1}^\top \end{bmatrix} \in \mathbb{C}^{n \times n\ell}.$$

Consider the compact SVD given by

$$V(v_j \otimes I_n) = X\Sigma\widehat{X}^* \text{ and } U(v_j \otimes I_n) = Y\widehat{\Sigma}\widehat{Y}^*.$$

Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $A - \lambda B$. Then the columns of X form an orthonormal basis of $N(A - \lambda_j B)$ and the columns of Y form an orthonormal basis of $N((A - \lambda_j B)^\top)$.

Proof. By Theorem 5.3.3 and (5.13), we have $N(A - \lambda_j B) = R(P_j) = R(V(v_j \otimes I_n))$ and $N((A - \lambda_j B)^\top) = R(Q_j^\top) = R(U(v_j \otimes I_n))$. Since $\operatorname{rank}(P_j) = \operatorname{rank}(Q_j) = m_j$, it follows from the compact SVD that $X \in \mathbb{C}^{n \times m_j}$ and $Y \in \mathbb{C}^{n \times m_j}$ are isometry and $R(X) = R(V(v_j \otimes I_n)) = N(A - \lambda_j B)$ and $R(Y) = R((U(v_j \otimes I_n))) = N((A - \lambda_j B)^\top)$. This completes the proof. □

Notice that in order to recover both left and right eigenvectors of $A - \lambda B$ corresponding to λ_j, we need compact SVD of $V(v_j \otimes I_n)$ and $U(v_j \otimes I_n)$. In contrast, for a standard value problem $Av = \lambda v$, a compact SVD of $V(v_j \otimes I_n)$ yields both right and left eigenvectors of A corresponding to λ_j (see, Theorem 5.2.10).

Remark 5.3.5. *Observe that if the curve Γ contains ℓ distinct eigenvalue of $A - \lambda B$ and each eigenvalue is semisimple then exactly ℓ matrix moments $S_0, \ldots, S_{\ell-1}$ are needed to recover the ℓ right spectral projections $P_1, \ldots, P_\ell$ of $A - \lambda B$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively. Similarly, the ℓ matrix moments $M_0, \ldots, M_{\ell-1}$ are needed to recover the ℓ left spectral projections $Q_1, \ldots, Q_\ell$ of $A - \lambda B$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively.*

Corollary 5.3.6. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $A - \lambda B$. Let P_j and Q_j be right and left spectral projections of $A - \lambda B$ corresponding to λ_j, respectively. Let v_j be given by $v_j = \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then for $j = 1 : \ell$, we have*

$$\begin{aligned} P_j &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zB - A)^{-1} B dz\right)/\alpha_j, \\ Q_j &= \left(\beta_1^j M_0 + \cdots + \beta_\ell^j M_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z) B (zB - A)^{-1} dz\right)/\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial and $\alpha_j := q_{v_j}(\lambda_j)$.

Proof. By Theorem 5.3.3, we have $P_j = V(v_j \otimes I_n)/\alpha_j$ and

$$\begin{aligned} V(v_j \otimes I_n) &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right) \\ &= \frac{1}{2\pi i}\int_\Gamma \sum_{k=1}^{\ell} \beta_k^j z^{k-1}(zB - A)^{-1} B dz \\ &= \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zB - A)^{-1} B dz, \end{aligned}$$

which yields the desired result.

Again by Theorem 5.3.3, we have $Q_j = (U(v_j \otimes I_n))^\top/\alpha_j$ and

$$\begin{aligned}(U(v_j \otimes I_n))^\top &= \left(\beta_1^j M_0 + \cdots + \beta_\ell^j M_{\ell-1}\right)\\ &= \frac{1}{2\pi i}\int_\Gamma \sum_{k=1}^{\ell} \beta_k^j z^{k-1} B(zB-A)^{-1} dz\\ &= \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z) B(zB-A)^{-1} dz,\end{aligned}$$

which yields the desired result.

□

5.4 Moment problems for bounded linear operators

Let $A \in L(X)$. Then a subspace $M \subset X$ is said to be invariant under A if $AM \subset M$. Let $P \in L(X)$ be a nontrivial projection, that is, $P^2 = P$ and $0 \neq P \neq I$. Set $Y := R(P)$ and $Z := N(P)$. Then relative to the decomposition $X = R(P) \oplus N(P)$ the operator A can be written as 2-by-2 operator matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} PAP & PA(I-P) \\ (I-P)AP & (I-P)A(I-P) \end{bmatrix}.$$

We have allowed a slight abuse of notation in the last equality above for simplicity of notation. Strictly speaking, $A_{11} = PA_{|R(P)}$ instead of PAP etc. Then

$$\begin{aligned} AY \subset Y &\iff A_{21} = 0 \iff (I-P)AP = 0 \iff AP = PAP\\ AZ \subset Z &\iff A_{12} = 0 \iff PA(I-P) = 0 \iff PA = PAP.\end{aligned}$$

Hence $AY \subset Y$ and $AZ \subset Z \iff AP = PA \iff A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$.

Suppose that $PA = AP$. Then $AY \subset Y$ and $AZ \subset Z$. The invariant subspaces Y and Z are called reducing subspaces of A. Set $A_Y := PAP$ and $A_Z := (I-P)A(I-P)$. Note that we consider A_Y and A_Z as the restrictions of A on Y and Z, respectively. Then

$$A = \begin{bmatrix} A_Y & 0 \\ 0 & A_Z \end{bmatrix} \text{ and } \sigma(A) = \sigma(A_Y) \cup \sigma(A_Z).$$

If $\sigma(A_Y) \cap \sigma(A_Z) = \emptyset$ then Y (resp., Z) is called the *spectral subspace* of A corresponding to $\sigma(A_Y)$ (resp., $\sigma(A_Z)$). The projection P is called the *spectral projection* of A corresponding to $\sigma(A_Y)$. Note that a spectral projection decomposes X and A as well as the spectrum $\sigma(A)$.

Remark 5.4.1. *Let $P \in L(X)$ be a projection and $\sigma_0 \subset \sigma(A)$ be compact. Then P is the spectral projection of A corresponding to σ_0 if and only if $PA = AP, \sigma(A_Y) = \sigma_0$ and $\sigma(A_Z) = \sigma(A) \setminus \sigma_0$ where $A_Y := PAP$ and $A_Z := (I-P)A(I-P)$ are restrictions of A on $Y := R(P)$ and $Z := N(P)$, respectively.*

Define $R(z) := (A - zI)^{-1}$ for $z \in \rho(A)$. Then $R(z)$ is called the *resolvent operator* of A. For $z, w \in \rho(A)$, we have the first resolvent identity

$$R(z) - R(w) = (z - w)R(z)R(w).$$

Consequently, the function $R : \rho(A) \longrightarrow L(X), z \longmapsto (A - zI)^{-1}$, is analytic.

Let $\Gamma \subset \rho(A)$ be a positively oriented rectifiable simple closed curve. For the rest of the thesis, we denote the region inside Γ by $\mathrm{Int}(\Gamma)$ and the region outside Γ by $\mathrm{Ext}(\Gamma)$. Thus $\mathbb{C} = \mathrm{Int}(\Gamma) \cup \Gamma \cup \mathrm{Ext}(\Gamma)$. Now, if f is analytic in an open set U and if $\sigma(A) \subset \mathrm{Int}(\Gamma) \subset U$ then, by functional calculus [23, p.14], we have

$$f(A) = \frac{-1}{2\pi i} \int_\Gamma f(z)R(z)dz. \tag{5.16}$$

Next, suppose that $\sigma_0 := \sigma(A) \cap \mathrm{Int}(\Gamma) \neq \emptyset$. Define

$$P := \frac{-1}{2\pi i} \int_\Gamma R(z)dz \quad \text{and} \quad S(z_0) := \frac{1}{2\pi i} \int_\Gamma \frac{R(z)}{z - z_0}dz \quad \text{for } z_0 \notin \Gamma.$$

Then P is the spectral projection of A corresponding to σ_0. In fact, the following result holds.

Theorem 5.4.2 (Spectral decomposition, [39, 33])**.** *Let $A \in L(X)$. Suppose that $\Gamma \subset \rho(A)$ and $\sigma(A) \cap \mathrm{Int}(\Gamma) \neq \emptyset$. Then the following results hold.*

(a) The operators A, P and $S(z_0)$ commute with each other.

(b) $P^2 = P$, that is, P is a projection.

(c) If $z_0 \in \mathrm{Int}(\Gamma)$ then $S(z_0)P = 0$ and $(A - z_0 I)S(z_0) = I - P$.

(d) If $z_0 \in \mathrm{Ext}(\Gamma)$ then $S(z_0)P = S(z_0)$ and $(A - z_0 I)S(z_0) = -P$.

(e) Set $Y := R(P)$ and $Z := N(P)$. Then $X = Y \oplus Z$ and $A = \begin{bmatrix} A_Y & 0 \\ 0 & A_Z \end{bmatrix}$, where $A_Y := PAP$ and $A_Z := (I-P)A(I-P)$. Further, $\Gamma \subset \rho(A_Y) \cap \rho(A_Z)$ and $\sigma(A) = \sigma(A_Y) \cup \sigma(A_Z)$. Furthermore, we have $\sigma(A_Y) = \sigma(A) \cap \mathrm{Int}(\Gamma)$ and $\sigma(A_Z) = \sigma(A) \cap \mathrm{Ext}(\Gamma)$.

Discrete eigenvalues. Let $\lambda \in \sigma(A)$ be an isolated spectral value. Suppose that $\sigma(A) \cap \mathrm{Int}(\Gamma) = \{\lambda\}$. Then

$$P := \frac{-1}{2\pi i}\int_{\Gamma} R(z)dz$$

is the spectral projection of A corresponding to λ. Thus $R(P)$ is the spectral subspace of A corresponding to λ. By Theorem 5.4.2, we have the spectral decomposition

$$A = \begin{bmatrix} A_Y & 0 \\ 0 & A_Z \end{bmatrix}, \quad \sigma(A_Y) = \{\lambda\} \text{ and } \sigma(A_Z) = \sigma(A) \setminus \{\lambda\}.$$

Definition 5.4.3. *Let $\lambda \in \sigma(A)$ be an isolated spectral value and P be the spectral projection of A corresponding to λ. If* $\mathrm{rank}(P) < \infty$ *then λ is called a* discrete eigenvalue *of A and $m := \mathrm{rank}(P)$ is called the* algebraic multiplicity *of λ. Set $\sigma_d(A) := \{\lambda \in \sigma(A) : \lambda$ is a discrete eigenvalue of $A\}$. Then $\sigma_d(A)$ is called the discrete spectrum of A.*

Let $\lambda \in \sigma_d(A)$. If $\mathrm{rank}(P) = m$ then there exists smallest $\ell \in \mathbb{N}$ such that $\ell \leq m$ and $R(P) = N((A - \lambda I)^{\ell})$. The number ℓ is called the ascent (or index) of λ. In fact, ℓ is the ascent of $\lambda \iff \lambda$ is a pole of $R(z)$ of order ℓ; see [39, p.92].

Definition 5.4.4. *Let $\lambda \in \sigma_d(A)$ and m be the algebraic multiplicity of λ. Then $g :=$* $\dim N(A - \lambda I)$ *is called the* geometric multiplicity *of λ. Obviously $g \leq m$. If $m = g = 1$ then λ is said to be a simple eigenvalue. On the other hand, λ is said to be* semisimple *(resp.,* defective*) if $m = g$ (resp., $g < m$).*

Obviously λ is a semisimple eigenvalue of $A \iff R(P) = N(A - \lambda I)$, where P is the spectral projection of A corresponding to λ.

Theorem 5.4.5 (Spectral theorem, [39, 33])**.** *Suppose that $\sigma_0 := \sigma(A) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_n\} \subset \sigma_d(A)$. Let P_j denote the spectral projection of A corresponding to λ_j for $j = 1 : n$. Then*

$$P = -\frac{1}{2\pi i}\int_{\Gamma} R(z)dz = P_1 + \cdots + P_n,$$

$P_iP_j = 0$ for $i \neq j$ and $R(P) = R(P_1) \oplus \cdots \oplus R(P_n)$. Further, we have

$$AP = \sum_{j=1}^{n}(\lambda_j P_j + D_j), \text{ where } D_j := (A - \lambda_j I)P_j$$

is a nilpotent operator for $j = 1 : n$. If ℓ_j is the ascent if λ_j then $D_j^{\ell_j} = 0$ for $j = 1 : n$.

In particular, if λ_j is semisimple for $j = 1 : n$, then $AP = \lambda_1 P_1 + \cdots + \lambda_n P_n$.

We now consider tracial moments of A. The next result establishes relation between tracial moments and discrete eigenvalues of A.

Theorem 5.4.6. *Suppose that $\sigma_0 := \sigma(A) \cap \text{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. Let m_j be the algebraic multiplicity of λ_j for $j = 1 : \ell$. Then for any analytic function f (analytic in an open set containing σ_0), we have*

$$s_f := \text{Tr}\left(\frac{-1}{2\pi i}\int_\Gamma f(z)\, R(z)dz\right) = m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell).$$

In particular, we have

$$\text{mean}(\sigma_0) := \frac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell} = \frac{\text{Tr}\left(\int_\Gamma z\, R(z)dz\right)}{\text{Tr}\left(\int_\Gamma R(z)dz\right)}.$$

Proof. Let P be the spectral projection of A corresponding to σ_0. Set $Y := R(P)$ and $Z := N(P)$. Then by Theorem 5.4.2, we have

$$A = \begin{bmatrix} A_Y & 0 \\ 0 & A_Z \end{bmatrix}, \ \sigma(A_Y) = \sigma_0 \text{ and } \sigma(A_Z) = \sigma(A) \setminus \sigma_0.$$

Note that $\dim(Y) = \text{rank}(P) = m_1 + \cdots + m_\ell$ and $A_Y \in BL(Y)$. Hence by (5.16) and the spectral mapping theorem, we have $\sigma(f(A_Y)) = f(\sigma_0) = \{f(\lambda_1), \ldots, f(\lambda_\ell)\}$ and $\text{Tr}(f(A_Y)) = m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell)$.

Since the function $w \longmapsto (A_Z - wI)^{-1}$ is analytic for $w \in \text{Int}(\Gamma) \cup \Gamma$, we have $\int_\Gamma f(w)(A_Z - wI)dw = 0$. By functional calculus (5.16), we have

$$\begin{aligned} s_f &= \text{Tr}\left(\frac{-1}{2\pi i}\int_\Gamma f(z)\, R(z)dz\right) = \text{Tr}\left(\frac{-1}{2\pi i}\int_\Gamma f(z)\begin{bmatrix} (A_Y - zI)^{-1} & 0 \\ 0 & (A_Z - zI)^{-1} \end{bmatrix} dz\right) \\ &= \text{Tr}\left(\begin{bmatrix} f(A_Y) & 0 \\ 0 & 0 \end{bmatrix}\right) = \text{Tr}(f(A_Y)) = m_1 f(\lambda_1) + \cdots + m_\ell f(\lambda_\ell). \end{aligned}$$

Finally, by considering $f(z) = 1$ and $f(z) = z$, we obtain the desired result for $\text{mean}(\sigma_0)$. This completes the proof. □

Now we define the p-th order *tracial moment* of A associated with σ_0 by

$$s_p := \text{Tr}\left(\frac{-1}{2\pi i}\int_\Gamma z^p\, R(z)dz\right), \quad p \in \mathbb{N} \cup \{0\}.$$

Then for $f_p(z) := z^p$, by Theorem 5.4.6, we have

$$s_p = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p, \quad p \in \mathbb{N} \cup \{0\}. \tag{5.17}$$

Consider the $p \times p$ Hankel matrix H_p and shifted Hankel matrix $\widehat{H}_p$ given by

$$H_p := \text{Hankel}(s_0, \ldots, s_{2p-2}) \text{ and } \widehat{H}_p := \text{Hankel}(s_1, \ldots, s_{2p-1}), \quad p \in \mathbb{N}.$$

We have the following result which is an analogue of Theorem 6.2.3.

Theorem 5.4.7. *Let $A \in L(X)$ and $\Gamma \subset \rho(A)$. Suppose that $\sigma(A) \cap \text{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. Let $m_1, \ldots, m_\ell$ be the algebraic multiplicities of the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively. Then the Hankel matrix H_ℓ is nonsingular. Further, $\text{eig}(\widehat{H}_\ell, H_\ell) = \{\lambda_1, \ldots, \lambda_\ell\}$ and $\lambda_1, ..., \lambda_\ell$ are simple eigenvalues of the $\ell \times \ell$ Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. The algebraic multiplicities $m_1, \ldots, m_\ell$ are given by the Vandermonde system*

$$\begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}. \tag{5.18}$$

The number ℓ satisfies $\ell = \text{rank}(H_\ell) = \text{rank}(H_m) = \text{rank}(H_{\ell+p})$ for any non-negative integer p, where $m := s_0 = m_1 + \cdots + m_\ell$.

Proof. A verbatim proof of Theorem 6.2.3 yields the desired results. □

Remark 5.4.8. *Observe that if there are ℓ distinct discrete eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A inside the curve Γ, that is, $\sigma_0 := \sigma(A) \cap \text{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$, then 2ℓ tracial moments $s_0, \ldots, s_{2\ell-1}$ of A are needed to recover the ℓ discrete eigenvalues.*

Theorem 6.3.2 provides a method for computing the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A from the tracial moments $s_0, \ldots, s_{2\ell-1}$. If the eigenvalues are real then by (5.17) the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ is real symmetric. In particular, if A is a self-adjoint operator on a Hilbert space then the pencil $\widehat{H}_\ell - \lambda H_\ell$ is real symmetric.

5.4.1 Operator moments and spectral projections

Let $A \in L(X)$. Let $A^* \in L(X^*)$ be the dual operator of A, where X^* is the dual space of X. As before, assume that $\sigma_0 := \sigma(A) \cap \text{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. By Theorem 6.3.2, all eigenvalues of $\widehat{H}_\ell - \lambda H_\ell$ are simple and $\text{eig}(\widehat{H}_\ell, H_\ell) = \sigma_0$. Consequently, an eigenvector v_j of $\widehat{H}_\ell - \lambda H_\ell$ corresponding to λ_j is unique up to a scalar multiple. Note that (λ_j, v_j) is also a left eigenpair of $\widehat{H}_\ell - \lambda H_\ell$.

Now consider the *operator moments*

$$S_p := \frac{1}{2\pi i}\int_\Gamma z^p (zI - A)^{-1} dz, \quad p = 0, 1, \ldots, \ell - 1. \tag{5.19}$$

Note that $P := S_0$ is the spectral projection of A corresponding to σ_0. We have the following result.

Theorem 5.4.9. *Let $A \in L(X)$. Suppose that $\Gamma \subset \rho(A)$ and $\sigma(A) \cap \mathrm{Int}(\Gamma) \neq \emptyset$. Set $P := S_0$. Then for any polynomial $q(z) \in \mathbb{C}[z]$, we have*

$$q(A)P = q(AP)P = \frac{-1}{2\pi i}\int_\Gamma q(z)R(z)dz.$$

In particular, if $v := \begin{bmatrix}\beta_1 & \cdots & \beta_\ell\end{bmatrix}^\top \in \mathbb{C}^\ell$ and $q(z) := \begin{bmatrix}1 & z & \cdots & z^{\ell-1}\end{bmatrix} v$ then

$$q(A)P = \beta_1 S_0 + \cdots + \beta_\ell S_{\ell-1}.$$

Proof. By Theorem 5.4.2, we have $AP = PA$. Hence $(AP)^m = A^m P$ for $m \in \mathbb{N} \cup \{0\}$. Consequently, we have $q(AP)P = q(A)P$ for $q(z) \in \mathbb{C}[z]$. Now using the fact that $AR(z) = (A - zI + zI)R(z) = I + zR(z)$ for $z \in \Gamma$, we have

$$\begin{aligned} AP &= \frac{-1}{2\pi i}\int_\Gamma AR(z)dz = \frac{-1}{2\pi i}\int_\Gamma [I + zR(z)]dz = \frac{-1}{2\pi i}\int_\Gamma zR(z)dz, \\ A^2 P &= \frac{-1}{2\pi i}\int_\Gamma zAR(z)dz = \frac{-1}{2\pi i}\int_\Gamma z[I + zR(z)]dz = \frac{-1}{2\pi i}\int_\Gamma z^2 R(z)dz. \end{aligned}$$

Hence the desired integral representation of $q(A)P$ follows by induction on m. The equality $q(A)P = \beta_1 S_0 + \cdots + \beta_\ell S_{\ell-1}$ follows from (5.19) and the integral representation of $q(A)P$. □

Remark 5.4.10. *Observe that $S_j = A^j P = PA^j$ for $j \in \mathbb{N} \cup \{0\}$. Hence $R(S_j) \subset R(P)$ for $j \in \mathbb{N} \cup \{0\}$. More generally, for $\beta_1, \ldots, \beta_\ell$ in $\mathbb{C}$, we have*

$$R(\beta_1 S_0 + \cdots + \beta_\ell S_{\ell-1}) \subset R(P).$$

For $x \in X$ and $\ell \in \mathbb{N}$, consider the Krylov subspace

$$\mathcal{K}_\ell(A, x) := \mathrm{span}(x, Ax, \ldots, A^{\ell-1}x).$$

Then for any $x \in X$ and $\beta_1, \ldots, \beta_\ell$ in $\mathbb{C}$, we have

$$(\beta_1 S_0 + \cdots + \beta_\ell S_{\ell-1})x \in \mathcal{K}_\ell(A, Px) \subset R(P).$$

Next, consider the operator moments of the dual operator A^* given by

$$S_p^* := \frac{1}{2\pi i}\int_\Gamma z^p(zI - A^*)^{-1}dz, \quad p = 0, 1, \ldots, \ell - 1.$$

Here we have used the generic notation I to denote the identity operator on the dual space X^*. Note that $P_0^* := S_0^*$ is the spectral projection of A^* corresponding to σ_0 and $R(P^*)$ is the spectral subspace of A^* corresponding to σ_0.

We now show that the spectral projection P_j of A corresponding to λ_j is an appropriate linear combination of the operator moments $S_0, \ldots, S_{\ell-1}$. We also show that a similar result holds for the spectral projection P_j^* of A^* corresponding to λ_j. For this purpose, we consider the operator matrices

$$\mathbf{V} := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \quad \text{and} \quad \mathbf{U}^* := \begin{bmatrix} S_0^* & \cdots & S_{\ell-1}^* \end{bmatrix}.$$

Then $\mathbf{V} : X^\ell \longrightarrow X$ and $\mathbf{U}^* : (X^*)^\ell \longrightarrow X^*$ are bounded linear operators defined as

$$\mathbf{V}\begin{bmatrix} x_1 \\ \vdots \\ x_\ell \end{bmatrix} := S_0x_1 + \cdots + S_{\ell-1}x_\ell \quad \text{and} \quad \mathbf{U}^*\begin{bmatrix} x_1^* \\ \vdots \\ x_\ell^* \end{bmatrix} := S_0^*x_1^* + \cdots + S_{\ell-1}^*x_\ell^*.$$

For $v := \begin{bmatrix} v_1 & \cdots & v_\ell \end{bmatrix}^\top \in \mathbb{C}^\ell, x \in X$ and $T \in L(X)$, define

$$v \otimes x := \begin{bmatrix} v_1x \\ \vdots \\ v_\ell x \end{bmatrix} \in X^\ell \text{ and } v \otimes T := \begin{bmatrix} v_1T \\ \vdots \\ v_\ell T \end{bmatrix} \in L(X, X^\ell).$$

More generally, if $B := [b_{ij}] \in \mathbb{C}^{m\times n}$ then $B \otimes T$ is the $m \times n$ operator matrix whose (i, j)-th entry is the operator $b_{ij}T$. Note that $(B \otimes T)(u \otimes x) = Bu \otimes Tx$ for $u \in \mathbb{C}^n$ and $x \in X$. Further, if $B \in \mathbb{C}^{n\times n}$ and $T \in L(X)$ are invertible then $B \otimes T$ is invertible and $(B \otimes T)^{-1} = B^{-1} \otimes T^{-1}$. Finally, for $x_1, \ldots, x_\ell$ in X, define $\begin{bmatrix} x_1 & \cdots & x_\ell \end{bmatrix} : \mathbb{C}^\ell \longrightarrow X$ by

$$\begin{bmatrix} x_1 & \cdots & x_\ell \end{bmatrix}\begin{bmatrix} v_1 \\ \vdots \\ v_\ell \end{bmatrix} := v_1x_1 + \cdots + v_\ell x_\ell.$$

The transformation $Q := \begin{bmatrix} x_1 & \cdots & x_\ell \end{bmatrix}$ is referred to as a quasi-matrix. Note that Q is an $n \times \ell$ matrix when $X = \mathbb{C}^n$.

Theorem 5.4.11. *Let* $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ *be eigenpairs of* $\widehat{H}_\ell - \lambda H_\ell$. *Define*

$$\mathbf{V} := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \in L(X^\ell, X) \text{ and } \mathbf{U}^* := \begin{bmatrix} S_0^* & \cdots & S_{\ell-1}^* \end{bmatrix} \in L((X^*)^\ell, X^*).$$

Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ *of* A *are semisimple. Let* I_X *and* I_{X^*} *denote the identity operators on* X *and* X^*, *respectively. Then for* $j = 1 : \ell$, *the spectral projections* P_j *and* P_j^* *of* A *and* A^*, *respectively, corresponding to* λ_j *are given by*

$$\begin{aligned} P_j &= \mathbf{V}(v_j \otimes I_X)/\alpha_j = \left(\frac{1}{2\pi i} \int_\Gamma q_{v_j}(z)(zI_X - A)^{-1} dz \right) /\alpha_j, \\ P_j^* &= \mathbf{U}^*(v_j \otimes I_X^*)/\alpha_j = \left(\frac{1}{2\pi i} \int_\Gamma q_{v_j}(z)(zI_{X^*} - A^*)^{-1} dz \right) /\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ *and* $\alpha_j := q_{v_j}(\lambda_j)$.

Let $x \in X$ *and* $x^* \in X^*$. *For* $j = 1 : \ell$, *define*

$$\begin{aligned} \mathbf{x}_j &:= \mathbf{V}(v_j \otimes x) = \begin{bmatrix} S_0 x \cdots S_{\ell-1} x \end{bmatrix} v_j \in X, \\ \mathbf{x}_j^* &:= \mathbf{U}(v_j \otimes x^*) = \begin{bmatrix} S_0^* x^* \cdots S_{\ell-1}^* x^* \end{bmatrix} v_j \in X^*. \end{aligned}$$

Then $\mathbf{x}_j$ *(resp.,* $\mathbf{x}_j^*$*) is an eigenvector of* A *(resp.,* A^**) corresponding to* λ_j *whenever* $\mathbf{x}_j \neq 0$ *(resp.,* $\mathbf{x}_j^* \neq 0$*).*

Proof. Since $\lambda_1, \ldots, \lambda_\ell$ are semisimple discrete eigenvalues of A, by Theorem 5.4.5, we have $P = P_1 + \cdots + P_\ell$ and $AP = \lambda_1 P_1 + \cdots + \lambda_\ell P_\ell$. Now by Theorem 5.4.9 and (5.19), we have $S_j = (AP)^j P = \lambda_1^j P_1 + \cdots + \lambda_\ell^j P_\ell$ for $j = 0, 1, \ldots, \ell - 1$. Note that

$$S_j = \lambda_1^j P_1 + \cdots + \lambda_\ell^j P_\ell = \begin{bmatrix} P_1 & \cdots & P_\ell \end{bmatrix} \begin{bmatrix} \lambda_1^j \otimes I_X \\ \vdots \\ \lambda_\ell^j \otimes I_X \end{bmatrix}, \; j = 0, 1, \ldots, \ell - 1.$$

This shows that $\mathbf{V} = \begin{bmatrix} P_1 & \cdots & P_\ell \end{bmatrix} (V_\ell^\top \otimes I_X)$ which in turn yields

$$P_j = \mathbf{V}(V_\ell^{-\top} e_j \otimes I_X),$$

where V_ℓ is the $\ell \times \ell$ Vandermonde matrix given in Theorem 6.2.3. Observe that $\alpha_j = \begin{bmatrix} 1 & \lambda_j & \cdots & \lambda_j^{\ell-1} \end{bmatrix} v_j = e_j^\top V_\ell^\top v_j$ for $j = 1 : \ell$. Hence by (6.5), we have

$$(\Lambda_\ell - \lambda_j I_\ell) V_\ell^\top v_j = 0 \Longrightarrow V_\ell^\top v_j = \alpha_j e_j \Longrightarrow V_\ell^{-\top} e_j = v_j/\alpha_j$$

for $j = 1 : \ell$. Hence we have $P_j = \mathbf{V}(v_j \otimes I_X)/\alpha_j$ for $j = 1 : \ell$. The integral representation of P_j follows from Theorem 5.4.9. The proof for P_j^* is similar.

Since $AP_j = \lambda_j P_j$, it follows that $P_j x$ is an eigenvector of A corresponding to λ_j provided that $P_j x \neq 0$. Hence $\alpha_j P_j x = \mathbf{V}(v_j \otimes x) = \begin{bmatrix} S_0 x \cdots S_{\ell-1} x \end{bmatrix} v_j = \mathbf{x}_j$ is an eigenvector of A corresponding to λ_j whenever $\mathbf{x}_j \neq 0$. Similarly, it follows that $\alpha P_j^* x^* = \mathbf{U}^*(v_j \otimes x^*) = \begin{bmatrix} S_0^* x^* \cdots S_{\ell-1}^* x^* \end{bmatrix} v_j = \mathbf{x}_j^*$ is a left eigenvector of A corresponding to λ_j whenever $\mathbf{x}_j^* = \alpha_j P_j^* x^* \neq 0$. □

Remark 5.4.12. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that the eigenvector v_j is normalized such that $q_{v_j}(\lambda_j) := \begin{bmatrix} 1 & \lambda_j & \cdots & \lambda_j^{\ell-1} \end{bmatrix} v_j = 1$ for $j = 1 : \ell$. Then we have*

$$P_j = \mathbf{V}(v_j \otimes I_X) = \frac{1}{2\pi i} \int_\Gamma q_{v_j}(z)(zI_X - A)^{-1} dz \; f\,or\; j = 1 : \ell.$$

Let the eigenvector v_j be given by $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then

$$P_j = \mathbf{V}(v_j \otimes I_X) = \beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}$$

shows that each P_j is a linear combination of $S_0, \ldots, S_{\ell-1}$ where the scalar coefficients are components of the normalized eigenvector v_j for $j = 1 : \ell$.

An alternative view point of the recovery of spectral projections is provided by interpolating polynomials. Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A are semisimple. Then by Theorem 5.4.5, we have $AP = \lambda_1 P_1 + \cdots + \lambda_\ell P_\ell$. Consequently,

$$q(A)P = q(AP)P = q(\lambda_1)P_1 + \cdots + q(\lambda_\ell)P_\ell$$

for any polynomial $q(z) \in \mathbb{C}[z]$. For $j = 1 : \ell$, let $q_j(z)$ be an interpolating polynomial of degree at most $\ell - 1$ such that $q_j(\lambda_i) = \delta_{ij}$ for $i = 1 : \ell$, where δ_{ij} is the Kronecker delta. Then we have $q_j(A)P = P_j$ for $j = 1 : \ell$. The next result gives a choice of the interpolating polynomials $q_1(z), \ldots, q_\ell(z)$.

Theorem 5.4.13. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\begin{bmatrix} 1 & \lambda_j & \cdots & \lambda_j^{\ell-1} \end{bmatrix} v_j = 1$ for $j = 1 : \ell$. Assume that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of A are semisimple. Consider the polynomial $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ for $j = 1 : \ell$. Then we have $q_{v_1}(z) + \cdots + q_{v_\ell}(z) = 1$ and $q_{v_j}(\lambda_i) = \delta_{ij}$ for $i = 1 : \ell$ and $j = 1 : \ell$. Further, for $j = 1 : \ell$, we have*

$$\begin{aligned} P_j &= q_{v_j}(A)P = \frac{1}{2\pi i} \int_\Gamma q_{v_j}(z)(zI_X - A)^{-1} dz \\ P_j^* &= q_{v_j}(A^*)P^* = \frac{1}{2\pi i} \int_\Gamma q_{v_j}(z)(zI_{X^*} - A^*)^{-1} dz. \end{aligned}$$

Proof. By (6.5), we have $(\widehat{H}_\ell - \lambda_j H_\ell)v_j = 0 \Longrightarrow (\Lambda_\ell - \lambda_j I_\ell)V_\ell^\top v_j = 0$ for $j = 1 : \ell$. This shows that for $j = 1 : \ell$, we have $V_\ell^\top v_j = \alpha_j e_j$ for some scalar α_j. We claim that $\alpha_j = 1$ for $j = 1 : \ell$. Indeed, $\alpha_j = e_j^\top V_\ell^\top v_j = \begin{bmatrix} 1 & \lambda_j & \cdots & \lambda_j^{\ell-1} \end{bmatrix} v_j = q_{v_j}(\lambda_j) = 1$ for $j = 1 : \ell$. Thus we have $V_\ell^\top v_j = e_j$ for $j = 1 : \ell$ which in turn shows that

$$q_{v_j}(\lambda_i) = \begin{bmatrix} 1 & \lambda_i & \cdots & \lambda_i^{\ell-1} \end{bmatrix} v_j = e_i^\top V_\ell^\top v_j = e_i^\top e_j = \delta_{ij}$$

for $i = 1 : \ell$ and $j = 1 : \ell$.

Next, define $q(z) := q_{v_1}(z) + \cdots + q_{v_\ell}(z)$. Note that $q(z)$ is a polynomial of degree at most $\ell - 1$ and that $q(\lambda_j) - 1 = 0$ for $j = 1 : \ell$. Hence by fundamental theorem of algebra we have $q(z) = 1$ for all $z \in \mathbb{C}$.

Finally, by Theorem 5.4.5 we have $AP = \lambda_1 P_1 + \cdots + \lambda_\ell P_\ell$ which in turn yields $q_{v_j}(AP)P = q_{v_j}(\lambda_1)P_1 + \cdots + q_{v_j}(\lambda_j)P_j + \cdots + q_{v_\ell}(\lambda_\ell)P_\ell = P_j$ for $j = 1 : \ell$. Now, by Theorem 5.4.9 we have $P_j = q_{v_j}(AP)P = q_{v_j}(A)P = \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI_X - A)^{-1}dz$ for $j = 1 : \ell$. The proof for P_j^* is similar. □

Remark 5.4.14. *The proof of Theorem 5.4.13 shows that* $V_\ell^\top \begin{bmatrix} v_1 & \cdots & v_\ell \end{bmatrix} = I_\ell$. *In other words, the normalized eigenvector matrix* $\begin{bmatrix} v_1 & \cdots & v_\ell \end{bmatrix} \in \mathbb{C}^{\ell\times\ell}$ *is the inverse of the Vandermonde matrix* $V_\ell^\top$. *If the eigenvectors* $v_1, \ldots, v_\ell$ *are not normalized then we have* $V_\ell^\top \begin{bmatrix} v_1 & \cdots & v_\ell \end{bmatrix} = \mathrm{diag}(q_{v_1}(\lambda_1), \ldots, q_{v_\ell}(\lambda_\ell))$.

Remark 5.4.15. *In view of Theorem 5.4.6 and (5.19), we have*

$$s_j = \mathrm{Tr}(S_j) = m_1\lambda_1^j + \cdots + m_\ell\lambda_\ell^j, \text{ for } j \in \mathbb{N}\cup\{0\}.$$

Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ *be eigenpairs of* $\widehat{H}_\ell - \lambda H_\ell$. *Let the eigenvector* v_j *be given by* $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$, $j = 1 : \ell$. *For* $j = 1 : \ell$, *consider the operators*

$$B_j := \mathbf{V}(v_j \otimes I_X) = \beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1} \in L(X).$$

If $\lambda_1, \ldots, \lambda_\ell$ *are semisimple eigenvalues of* A *then we have seen that* $B_j = q_{v_j}(\lambda_j)P_j$ *for* $j = 1 : \ell$. *Note that* $B_1, \ldots, B_\ell$ *are finite rank operators. Hence, if* X *is a Hilbert space then, as in Theorem 5.2.10, singular value decomposition of the operators* $B_1, \ldots, B_\ell$ *yield orthonormal bases of* $N(A - \lambda_1 I), \ldots, N(A - \lambda_\ell I)$, *respectively.*

The integral representation of spectral projections in Theorem 5.4.13 has a filter function interpretation. Consider the function $\phi : \mathbb{C} \longrightarrow \mathbb{C}$ given by

$$\phi(w) := \frac{1}{2\pi i}\int_\Gamma \frac{dz}{z - w}.$$

Then $\phi(w) = 1$ when $w \in \sigma(A) \cap \mathrm{Int}(\Gamma)$ and $\phi(w) = 0$ when $w \in \mathrm{Ext}(\Gamma)$. Thus $\phi(w)$ acts as a filter function which filters out the spectral values in $\sigma(A) \cap \mathrm{Ext}(\Gamma)$. Next, observe that

$$P = \phi(A) = \frac{1}{2\pi i}\int_\Gamma (zI - A)^{-1} dz$$

is the spectral projection of A corresponding to $\sigma_0 := \sigma(A) \cap \mathrm{Int}(\Gamma)$.

Let $q_{v_1}(z), \ldots, q_{v_\ell}(z)$ be as in Theorem 5.4.13. For $j = 1 : \ell$, define $\phi_j : \mathbb{C} \longrightarrow \mathbb{C}$ by

$$\phi_j(w) := \frac{1}{2\pi i}\int_\Gamma \frac{q_{v_j}(z)}{z - w} dz.$$

Then $\phi_j(w) = 0$ for $w \in \mathrm{Ext}(\Gamma)$ and $q_{v_j}(\lambda_i) = \delta_{ij}$ for $i = 1 : \ell$ and $j = 1 : \ell$. Thus $q_{v_j}(z)$ acts as a filter function which filters out the spectral values in $\sigma(A) \setminus \{\lambda_j\}$ for $j = 1 : \ell$. Observe that

$$P_j = \phi_j(A) = \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)(zI - A)^{-1} dz$$

is the spectral projection of A corresponding to λ_j for $j = 1 : \ell$.

Since $q_{v_1}(z) + \cdots + q_{v_\ell}(z) = 1$, it follows that $\phi(z) = \phi_1(z) + \cdots + \phi_\ell(z)$. The filter functions $\phi(z)$ and $\phi_1(z), \ldots, \phi_\ell(z)$ yield approximate filter functions when the contour integrals are approximated by a numerical quadrature. Indeed, for the nodes $z_1, \ldots, z_N$ and weights $w_1, \ldots, w_N$, we have

$$\begin{aligned}
\phi(z) &= \frac{1}{2\pi i}\int_\Gamma \frac{d\xi}{\xi - z} \approx \sum_{k=1}^{N} \frac{w_k}{z_k - z} =: \widehat{\phi}(z) \\
\phi_j(z) &= \frac{1}{2\pi i}\int_\Gamma \frac{q_{v_j}(\xi)}{\xi - z} d\xi \approx \sum_{k=1}^{N} w_k \frac{q_{v_j}(z_k)}{z_k - z} =: \widehat{\phi}_j(z), \ j = 1 : \ell.
\end{aligned}$$

Since $q_{v_1}(z) + \cdots + q_{v_\ell}(z) = 1$, it follows that $\widehat{\phi}(z) = \widehat{\phi}_1(z) + \cdots + \widehat{\phi}_\ell(z)$. We have $P \approx \widehat{\phi}(A)$ and $P_j \approx \widehat{\phi}_j(A)$ for $j = 1 : \ell$. In other words, $\widehat{\phi}(A)$ and $\widehat{\phi}_j(A), j = 1 : \ell$, are approximate spectral projections.

We mention that our solutions to the spectral recovery problems are dependent on the solution of a generalized Hankel eigenvalue problem. As such, it is well known that numerical problem involving Hankel matrices tend to be ill-conditioned especially when the size of the Hankel matrix is large. Hence it is advisable to compute a small number of eigenvalues of A using the spectral recovery method. Note that the size of the Hankel pencil is independent of the multiplicities of the eigenvalues of A inside the curve Γ.

For the recovery of the spectral projections $P_1, \ldots, P_\ell$ from the operator moments $S_0, \ldots, S_{\ell-1}$ it was necessary to assume in Theorem 5.4.11 that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of A. This facilitated recovery of $P_1, \ldots, P_\ell$ from the eigenvectors of the

Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. On the other hand, if we drop the assumption that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of A then, as the following result shows, it is still possible to recover $P_1, \ldots, P_\ell$ from the operator moments $S_0, \ldots, S_N$ for some $N \in \mathbb{N}$. Indeed, we have $P_j = \beta_0^j S_0 + \cdots + \beta_N^j S_N$ but it is not known how to determine N and the scalars $\beta_0^j, \ldots, \beta_N^j$.

Theorem 5.4.16. *Let $A \in L(X)$ and $\tau := \sigma(A) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(A)$. Then there exist polynomials $p_1(z), \ldots, p_\ell(z)$ such that $p_j(\lambda_i) = \delta_{ij}$ and $p_1(z)+\cdots+p_\ell(z) = 1$. The spectral projection of A corresponding to λ_j is given by*

$$P_j = \frac{1}{2\pi i}\int_\Gamma p_j(z)(zI - A)^{-1} dz \quad \text{for } j = 1 : \ell.$$

Suppose that $p_j(z)$ is given by $p_j(z) = \beta_0^j + \beta_1^j z + \cdots + \beta_{N_j}^j z^{N_j}$. Then we have

$$P_j = \beta_0^j S_0 + \cdots + \beta_{N_j}^j S_{N_j} \text{ for } j = 1 : \ell.$$

Proof. Let $m(z) := \prod_{j=1}^{\ell}(z - \lambda_j)^{\nu_j}$ be the minimal polynomial of AP, where P be the spectral projection of A corresponding to τ. Set

$$\phi_j(z) := (z - \lambda_j)^{\nu_j} \text{ and } q_j(z) := m(z)/\phi_j(z) \text{ for } j = 1 : \ell.$$

For $i \neq j$, we have $q_i(z)q_j(z) = m(z)r(z)$ for some $r(z) \in \mathbb{C}[z]$. Consequently, we have $q_i(AP)q_j(AP) = 0$ for $i \neq j$. Since $q_1(z) \ \ldots, q_\ell(z)$ are coprime, there exist $a_1(z), \ldots, a_\ell(z)$ in $\mathbb{C}[z]$ such that $a_1(z)q_1(z)+\cdots+a_\ell(z)q_\ell(z) = 1$. Define $p_j(z) := a_j(z)q_j(z)$ for $j = 1 : \ell$. Then it follows that $p_j(\lambda_i) = \delta_{ij}$ and $p_1(z) + \cdots + p_\ell(z) = 1$, where δ_{ij} is Kronecker delta. Hence by Theorem 5.4.9, we have

$$p_1(A)P + \cdots + p_\ell(A)P = P. \tag{5.20}$$

Define $P_j := p_j(A)P$ for $j = 1 : \ell$. Then $AP_j = P_jA$ for $j = 1 : \ell$. Further, we have

$$P_iP_j = p_i(A)p_j(A)P = a_i(A)q_i(A)a_j(A)q_j(A)P = a_i(A)a_j(A)q_i(AP)q_j(AP)P = 0$$

for $i \neq j$. By (5.20) we have $P_j^2 = P_j$ for $j = 1 : \ell$. Hence $R(P) = R(P_1) \oplus \cdots \oplus R(P_\ell)$.

We now show P_i is the spectral projection of A corresponding to λ_i by showing that $R(P_i) = N((A - \lambda_i I)^{\nu_i})$. Let $u \in N((A - \lambda_i I)^{\nu_i})$. Then $Pu = u$ and $\phi_i(AP)u = 0$ which implies that $q_j(AP)u = 0$ for $j \neq i$. Hence $P_ju = 0$ for $j \neq i$. By (5.20) we have $u = Pu = P_1u + \cdots + P_iu + \cdots + P_\ell u = P_iu \implies u \in R(P_i)$. Conversely, let

$u \in R(P_i)$. Then $u = P_i u = a_i(AP)q_i(AP)u = q_i(A)w$, where $w := a_i(AP)u = a_i(A)u$. Now $\phi_i(A)u = \phi_i(A)q_i(A)w = m(A)w = m(AP)w = 0 \Longrightarrow u \in N(\phi_i(A))$.

Finally, by Theorem 5.4.9, we have

$$\begin{aligned} P_j &= p_j(A)P = \frac{1}{2\pi i}\int_\Gamma p_j(z)(zI - A)^{-1}dz \\ &= \beta_0^j S_0 + \cdots + \beta_{N_j}^j S_{N_j} \quad \text{for} \quad j = 1 : \ell. \end{aligned}$$

□

To sum up: We have shown that the moment problem for discrete eigenvalues has an elegant solution that utilizes Hankel matrices. When the discrete eigenvalues are semisimple, we have shown that the moment problem for the associated spectral projections also has an elegant solution. We have also shown that the individual spectral projections admit an integral representation that involves interpolating polynomials and that such interpolating polynomials are readily given by eigenvectors of the Hankel pencils. The integral representation provides a numerical quadrature based algorithm for computing the spectral projections. Also, we have constructed filter functions for individual spectral projections by utilizing the interpolating polynomials given by eigenvectors of the Hankel pencil.

6

Moment problems for holomorphic matrices

Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular. Let $\lambda_1, \ldots, \lambda_\ell$ be eigenvalues of $\mathbf{T}(z)$ inside the region enclosed by a curve Γ. The main objective of this chapter is to determine $\lambda_1, \ldots, \lambda_\ell$, their algebraic multiplicities, and their corresponding eigenvectors from the moments

$$s_p := \frac{1}{2\pi i}\int_\Gamma z^p \mathrm{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz \text{ and } S_p := \frac{1}{2\pi i}\int_\Gamma z^p \mathbf{T}(z)^{-1}dz \tag{6.1}$$

for $p \in \mathbb{Z}_+$. Thus, approximating the contour integrals by numerical quadrature, we obtain a numerical method for solving the nonlinear eigenvalue problem $\mathbf{T}(\lambda)v = 0$.

6.1 Introduction

Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular. Consider the nonlinear eigenvalue problem:

$$\text{Find } (\lambda, v) \in \mathbb{C}\times\mathbb{C}^n \text{ with } v \neq 0 \text{ such that } \mathbf{T}(\lambda)v = 0.$$

Set $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$. Then $\mathbf{M}(z)$ is meromorphic and λ is a pole of $\mathbf{M}(z) \iff \lambda$ is an eigenvalue of $\mathbf{T}(z)$. Set $\mathcal{O} := \mathrm{Int}(\Gamma)$. Then $\wp_{\mathcal{O}}(\mathbf{M}) = \{\lambda_1, \ldots, \lambda_\ell\}$. We have already seen in Chapter 4 that the poles of $\mathbf{M}(z)$ in $\mathcal{O}$ can be computed from a local minimal realization

$$\mathbf{M}(z) \simeq_{\mathcal{O}} C(zI_r - A)^{-1}B.$$

In such a case, $\mathrm{eig}(A) = \{\lambda_1, \ldots, \lambda_\ell\}$ and if v_i is an eigenvector of A corresponding to λ_i then by Theorem 5.2.10 Cv_i is an eigenvector of $\mathbf{T}(z)$ corresponding to λ_i. The matrices (C, A, B) can be computed from the Markov parameters (matrix moments) S_p as given in (6.1) for $p \in \mathbb{Z}_+$. Indeed, (C, A, B) can be extracted from the block Hankel matrix

$\mathbb{H}_p$ and the shifted block Hankel matrix $\widehat{\mathbb{H}}_p$ given by

$$\mathbb{H}_p := \begin{bmatrix} S_0 & \cdots & S_{p-1} \\ \vdots & \ddots & \vdots \\ S_{p-1} & \cdots & S_{2p-2} \end{bmatrix} \text{ and } \widehat{\mathbb{H}}_p := \begin{bmatrix} S_1 & \cdots & S_p \\ \vdots & \ddots & \vdots \\ S_p & \cdots & S_{2p-1} \end{bmatrix} \text{ for } p \in \mathbb{Z}_+$$

by employing one of the algorithms outlined in Chapter 4.

The main objective of this chapter is to present an alternative method. As in Chapter 5, we consider the tracial moments s_p as defined in (6.1) and solve the following moment problem.

Problem-C: *Given the tracial moments $s_0, \ldots, s_p$, for an appropriate $p \in \mathbb{N}$, determine the ℓ distinct eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ and their algebraic multiplicities.*

We show that exactly 2ℓ moments, namely, $s_0, \ldots, s_{2\ell-1}$ are needed in order to determine the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ and their algebraic multiplicities. As in the case of a linear operator, we show that $\lambda_1, \ldots, \lambda_\ell$ are eigenvalues of the $\ell \times \ell$ Hankel matrix pencil $\widehat{H}_\ell - \lambda H_\ell$ irrespective of the multiplicities of the eigenvalues, where

$$H_\ell := \begin{bmatrix} s_0 & \cdots & s_{\ell-1} \\ \vdots & \ddots & \vdots \\ s_{\ell-1} & \cdots & s_{2\ell-2} \end{bmatrix} \text{ and } \widehat{H}_\ell := \begin{bmatrix} s_1 & \cdots & s_\ell \\ \vdots & \ddots & \vdots \\ s_\ell & \cdots & s_{2\ell-1} \end{bmatrix}.$$

Now consider $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$. Then $\mathbf{M}(z)$ is meromorphic in the domain $\mathcal{O}$ and the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ are the poles of $\mathbf{M}(z)$ in $\mathcal{O}$. Let $\mathbf{P}_j := \mathrm{Res}(\lambda_j, \mathbf{M})$ be the residue of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. Consider the matrix moments

$$S_p := \frac{1}{2\pi i}\int_\Gamma z^p \mathbf{T}(z)^{-1} dz = \frac{1}{2\pi i}\int_\Gamma z^p \mathbf{M}(z) dz \quad \text{for} \quad p \in \mathbb{Z}_+.$$

We show that $S_0 = \mathbf{P}_1 + \cdots + \mathbf{P}_\ell$. Further, $R(\mathbf{P}_j) = N(\mathbf{T}(\lambda_j))$ when λ_j is a semisimple eigenvalue of $\mathbf{T}(z)$ for $j = 1 : \ell$. We consider the following matrix moment problem.

Problem-D: *Given the matrix moments $S_0, \ldots, S_p$, for an appropriate $p \in \mathbb{N}$, determine the residues $\mathbf{P}_1, \ldots, \mathbf{P}_\ell$ of $\mathbf{T}(z)$ at the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively.*

We show that exactly ℓ matrix moments $S_0, \ldots, S_{\ell-1}$ are needed to determine the residues $\mathbf{P}_1, \ldots, \mathbf{P}_\ell$ when all the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ are semisimple.

We also consider analogues of Problem-C and Problem-D when $\mathbf{T}(z)$ is a holomorphic operator-valued function and $\lambda_1, \ldots, \lambda_\ell$ are discrete eigenvalues of $\mathbf{T}(z)$. We provide two different proofs - one for matrix-valued functions and another for operator-valued functions - highlighting the differences between finite and infinite dimensional problems.

6.2 Moment problems for holomorphic matrices

Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular and $\Gamma \subset \rho_\Omega(\mathbf{T})$ be a positively oriented rectifiable simple closed curve such that $\sigma_\Omega(\mathbf{T}) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\}$. Set $\mathcal{O} := \mathrm{Int}(\Gamma)$. We assume that $\mathcal{O} \subset \Omega$. Then $\mathcal{O}$ is simply connected and $\sigma_{\mathcal{O}}(\mathbf{T}) = \{\lambda_1, \ldots, \lambda_\ell\}$. Let $m_1, \ldots, m_\ell$ be the algebraic multiplicities of $\lambda_1, \ldots, \lambda_\ell$, respectively.

Let $f \in \mathbb{H}(\Omega)$. Let $z_1, \ldots, z_\ell$ be distinct zeros of $f(z)$ in $\mathcal{O}$ with multiplicities $n_1, \ldots, n_\ell$, respectively. Then $N := n_1 + \cdots + n_\ell$ is the total multiplicity of the zeros and by the argument principle [58], we have

$$\frac{1}{2\pi i}\int_\Gamma \frac{f'(z)}{f(z)}dz = n_1 + \cdots + n_\ell = N.$$

Further, if $g \in \mathbb{H}(\Omega)$ then by generalized argument principle [58], we have

$$\frac{1}{2\pi i}\int_\Gamma \frac{g(z)f'(z)}{f(z)}dz = n_1 g(z_1) + \cdots + n_\ell g(z_\ell). \tag{6.2}$$

Theorem 6.2.1. *Let* $\mathrm{mean}(\lambda_1, \ldots, \lambda_\ell) := \dfrac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell}$ *be the arithmetic mean of the eigenvalues* $\lambda_1, \ldots, \lambda_\ell$ *of* $\mathbf{T}(z)$*. Let* $g \in \mathbb{H}(\Omega)$*. Then we have*

$$\frac{1}{2\pi i}\int_\Gamma g(z)\mathrm{Tr}(\mathbf{T}^{-1}(z)\mathbf{T}'(z))dz = m_1 g(\lambda_1) + \cdots + m_\ell g(\lambda_\ell). \tag{6.3}$$

In particular, we have

$$\mathrm{mean}(\lambda_1, \ldots, \lambda_\ell) = \frac{\int_\Gamma z\,\mathrm{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz}{\int_\Gamma \mathrm{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz}.$$

Proof. Set $f(z) := \det(\mathbf{T}(z))$. Then, by the Jacobi formula, $f'(z) = \mathrm{Tr}(\mathrm{adj}(\mathbf{T}(z))\mathbf{T}'(z))$, where $\mathrm{adj}(A)$ is the adjugate of the matrix A. Here $f'(z)$ and $\mathbf{T}'(z)$ denote the derivative of $f(z)$ and $\mathbf{T}(z)$, respectively. Since $\mathrm{adj}(\mathbf{T}(z)) = \mathbf{T}(z)^{-1}\det(\mathbf{T}(z)) = \mathbf{T}(z)^{-1}f(z)$, we have $\mathrm{Tr}(\mathbf{T}^{-1}(z)\mathbf{T}'(z)) = f'(z)/f(z)$. Consequently, by (6.2) we have

$$\frac{1}{2\pi i}\int_\Gamma g(z)\mathrm{Tr}(\mathbf{T}^{-1}(z)\mathbf{T}'(z))dz = \frac{1}{2\pi i}\int_\Gamma \frac{g(z)f'(z)}{f(z)}dz = m_1 g(\lambda_1) + \cdots + m_\ell g(\lambda_\ell).$$

Taking $g(z) = 1$ and $g(z) = z$, we have

$$\begin{aligned} \frac{1}{2\pi i}\int_\Gamma \operatorname{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz &= m_1 + \cdots + m_\ell, \\ \frac{1}{2\pi i}\int_\Gamma z \operatorname{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz &= m_1\lambda_1 + \cdots + m_\ell\lambda_\ell \end{aligned}$$

which yield mean$(\lambda_1, \ldots, \lambda_\ell)$. □

Remark 6.2.2. *Let $\mathbf{M} \in \mathbb{M}(\Omega)^{n\times n}$ be regular. Suppose that $\sigma_{\mathcal{O}}(\mathbf{M}) = \{z_1, \ldots, z_\ell\}$ with total multiplicity* n *and $\wp_{\mathcal{O}}(\mathbf{M}) = \{\mu_1, \ldots, \mu_r\}$ with total multiplicity* p*. Then by a proof similar to that of Theorem 6.2.1 with $g(z) = 1$, we have*

$$\frac{1}{2\pi i}\int_\Gamma \operatorname{Tr}(\mathbf{M}(z)^{-1}\mathbf{M}'(z))dz = \ell - \mathrm{r}.$$

By Theorem 6.2.1, the p-th order tracial moment s_p is given by

$$s_p = \frac{1}{2\pi i}\int_\Gamma z^p \operatorname{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p \ \text{ for } \ p \in \mathbb{Z}_+. \tag{6.4}$$

As before, consider the Hankel matrix H_k, shifted Hankel matrix $\widehat{H}_k$, and Vandermonde matrix V_k for $k \in \mathbb{N}$. Note that V_ℓ is invertible and $\operatorname{rank}(V_k) = \ell$ for all $k \geq \ell$.

Theorem 6.2.3. *Let $\sigma_{\mathcal{O}}(\mathbf{T}) = \{\lambda_1, \ldots, \lambda_\ell\}$ and $m_1, \ldots, m_\ell$ be multiplicities of $\lambda_1, \ldots, \lambda_\ell$, respectively. Then $\lambda_1, ..., \lambda_\ell$ are the simple eigenvalues of the $\ell\times\ell$ Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. The algebraic multiplicities of the eigenvalues are given by the $\ell\times\ell$ Vandermonde system*

$$V_\ell \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}.$$

The number ℓ satisfies $\ell = \operatorname{rank}(H_\ell) = \operatorname{rank}(H_N) = \operatorname{rank}(H_{\ell+p})$ for any non-negative integer p, where $N := s_0 = m_1 + \cdots + m_\ell$.

Proof. Define $D_\ell := \operatorname{diag}(m_1, \ldots, m_\ell)$ and $\Lambda_\ell := \operatorname{diag}(\lambda_1, \ldots, \lambda_\ell)$. By (6.4), we have $s_p = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p$ for $p \in \mathbb{Z}_+$ which gives $V_\ell \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}.$

Again by (6.4), it follows that $H_k = V_k D_\ell V_k^\top$ and $\widehat{H}_k = V_k D_\ell \Lambda_\ell V_k^\top$ for $k \in \mathbb{N}$. Thus we have $H_\ell = V_\ell D_\ell V_\ell^\top$ and $\widehat{H}_\ell = V_\ell D_\ell \Lambda_\ell V_\ell^\top$. Since V_ℓ and D_ℓ are invertible,

$$\widehat{H}_\ell - \lambda H_\ell = V_\ell D_\ell (\Lambda_\ell - \lambda I_\ell) V_\ell^\top \tag{6.5}$$

shows that $\lambda_1, \ldots, \lambda_\ell$ are simple eigenvalues of the pencil $\widehat{H}_\ell - \lambda H_\ell$.

Finally, since $H_k = V_k D_\ell V_k^\top$ and $\text{rank}(V_k) = \text{rank}(V_\ell) = \ell$ for all $k \geq \ell$, it follows that $\text{rank}(H_{\ell+p}) = \text{rank}(H_N) = \text{rank}(H_\ell) = \ell$ for all non-negative integer p. □

Theorem 6.2.3 provides a method for computing the eigenvalues of $\mathbf{T}(z)$ contained in $\mathcal{O}$. The first step is to compute the trace-moment s_0, which gives the total number of eigenvalues (counting multiplicity) of $\mathbf{T}(z)$ in $\mathcal{O}$. The second step is to compute $\ell = \text{rank}(H_{s_0})$, which gives the number of distinct eigenvalues of $\mathbf{T}(z)$ in $\mathcal{O}$. Finally, solve the $\ell \times \ell$ Hankel eigenvalue problem $(\widehat{H}_\ell - \lambda H_\ell)v = 0$.

Remark 6.2.4. *Several comments on computation of the moments s_p are in order. The tracial moments s_p can be computed in three different ways giving rise to three mathematically equivalent methods for computing eigenvalues of $\mathbf{T}(z)$ contained in $\mathcal{O}$. Indeed, we have*

$$s_p = \frac{1}{2\pi i}\int_\Gamma z^p \, \text{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))dz \tag{6.6}$$

$$= \frac{1}{2\pi i}\int_\Gamma z^p \, \frac{f'(z)}{f(z)}dz \tag{6.7}$$

$$= \frac{1}{2\pi i}\int_\Gamma z^p \, \frac{d}{dz}(\log f(z))dz, \tag{6.8}$$

where $f(z) := \det(\mathbf{T}(z))$. However, approximation of $s_p \approx \sum_{j=0}^{N} w_j z_j^p \text{Tr}(\mathbf{T}(z_j)^{-1}\mathbf{T}'(z_j))$ using the integral (6.6) has several advantages over the integrals (6.7) and (6.8). Firstly, for the special case when $\mathbf{T}(z) = A - zI$ or $\mathbf{T}(z) = A - zB$, we have $\mathbf{T}'(z) = -I$ and $\mathbf{T}'(z) = -B$. Hence the integral in (6.6) is effectively derivative free. By contrast, the integral in (6.7) still requires the derivative $f'(z)$ which may be difficult to compute. Secondly, $\mathbf{T}(z)$ can always be expressed as $\mathbf{T}(z) = \phi_1(z)A_1 + \cdots + \phi_r(z)A_r$ for some analytic functions $\phi_1(z), \ldots, \phi_r(z)$. Hence the derivative $\mathbf{T}'(z)$ can be computed easily from the derivatives of $\phi_1(z), \ldots, \phi_r(z)$. This is certainly the case when $\mathbf{T}(z)$ is a matrix polynomial given by $\mathbf{T}(z) := A_0 + zA_1 + \cdots + z^r A_r$. Also, the MATLAB *toolbox [10] provides $\mathbf{T}'(z)$ for its collection of NEPs. Note that the derivatives of $\phi_1(z), \ldots, \phi_n(z)$ cannot be utilized for computing $f'(z)$. Thirdly, the integral in (6.6) requires solution of the linear system $\mathbf{T}(z)X = \mathbf{T}'(z)$ which, when restricted to a lower dimensional subspace, opens* *up the possibility of a projection based hybrid method. Fourth, the trace in (6.6) can be estimated as*

$$\mathrm{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z)) \approx \frac{1}{d}\sum_{j=1}^{d} v_j^\top \mathbf{T}(z)^{-1}\mathbf{T}'(z)v_j,$$

where $d < n$ and $v_1, \ldots, v_d$ are orthonormal sample vectors. Fifth, $\mathrm{Tr}(\mathbf{T}(z)^{-1}\mathbf{T}'(z))$ can also be approximated by solving the generalized eigenvalue problem

$$(\mathbf{T}(z) - \lambda\mathbf{T}'(z))v = 0$$

by a projection based method when the size of the matrix $\mathbf{T}(z)$ is large.

Notice that computation of $s_p \approx \sum_{j=0}^{N} w_j z_j^p \frac{f(z_j)}{f'(z_j)}$ via (6.7) reduces the problem of computing eigenvalues of $\mathbf{T}(z)$ to the problem of computing zeros of the analytic function $f(z)$. Thus, contour integration based algorithms developed for computing zeros of analytic functions can be utilized for solving NEPs, see [36, 17, 8] and the references therein. However, as remarked in [8], whether finding zeros of the determinant $f(z)$ can be an attractive option in solving large scale eigenvalue problems is yet to be seen.

Finally, computation of s_p via (6.8) has an advantage that it does not explicitly require $f'(z)$ as $f'(z)/f(z)$ is replaced with the derivative of $\log f(z)$ and hence integration by parts can be employed. The difficulty, however, is that $\log z$ is multi-valued and has a branch point at the origin. Therefore, one has to keep track of appropriate sheet on which $\log f(z)$ lies as z moves along the curve Γ [17]. Alternately, the integral (6.8) can be reformulated that is free from multivaluedness [31, 36]

$$s_p = \frac{-p}{2\pi i}\int_\Gamma z^{p-1}\ \log[(z-a)^{-s_0}f(z)]dz + s_0 a^p,$$

where $s_0 := \frac{1}{2\pi}[\arg f(z)]_\Gamma$ is total number of zeros of $f(z)$ inside the curve Γ and a is an arbitrary point inside Γ. In this case, the function $\log[(z-a)^{-s_0}f(z)]$ is single-valued unlike $\log f(z)$. We consider only the trace-moment s_p as given in (6.6).

Now approximating the tracial moments s_p by numerical quadrature

$$s_p \approx \sum_{j=1}^{N} w_j z_j^p \mathrm{Tr}(\mathbf{T}(z_j)^{-1}\mathbf{T}'(z))$$

we have the following method for computing eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$.

Algorithm for NEP

Input: Nodes z_j and weights w_j for j= $1 : N$.

Output: Eigenpair (λ_j, v_j) of $\hat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$.

- Compute $\mathtt{s_0} = \mathtt{sum}_{\mathtt{j=0}}^{\mathtt{N}} \mathtt{w_j} \,\mathrm{Tr}(\mathtt{T}(\mathtt{z_j})^{-1}\mathtt{T}'(\mathtt{z_j}))$
- Compute $\mathtt{s_p} := \mathtt{sum}_{\mathtt{j=0}}^{\mathtt{N}} \mathtt{w_j}\, \mathtt{z_j^p} \,\mathrm{Tr}(\mathtt{T}(\mathtt{z_j})^{-1}\mathtt{T}'(\mathtt{z_j}))$, $\mathtt{p} = 1 : 2\mathtt{s_0} - 2$
- Compute the Hankel matrix $\mathtt{H}_{\mathtt{s_0}}$ and compute $\ell = \mathrm{rank}(\mathtt{H}_{\mathtt{s_0}})$
- Compute the Hankel pencil $\widehat{\mathtt{H}}_\ell - \lambda \mathtt{H}_\ell$
- Compute eigenvalues and eigenvectors of $\widehat{\mathtt{H}}_\ell - \lambda \mathtt{H}_\ell$

We now describe the recovery of eigenvectors of $\mathbf{T}(z)$ from those of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. Recall that $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$ and $\mathbf{P}_j := \mathrm{Res}(\lambda_j, \mathbf{M}(z))$ is the residue of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. Let $\mathbf{M}(z) \simeq_{\mathcal{O}} C(zI_r - A)^{-1}B$ be a local minimal realization. Then

$$\mathbf{P}_j = \frac{1}{2\pi i}\int_{\Gamma_j} \mathbf{M}(z)dz = \frac{1}{2\pi i}\int_{\Gamma_j} C(zI_r - A)^{-1}Bdz = CP(\lambda_j, A)B, \tag{6.9}$$

where $P(\lambda_j, A)$ is the spectral projection of A corresponding to λ_j and Γ_j is a curve in $\mathcal{O}$ isolating λ_j from the rest of the poles of $\mathbf{M}(z)$ for $j = 1 : \ell$. Then

$$S_0 = \frac{1}{2\pi i}\int_{\Gamma} \mathbf{M}(z)dz = \sum_{j=1}^{\ell} \frac{1}{2\pi i}\int_{\Gamma_j} \mathbf{M}(z)dz = \mathbf{P}_1 + \cdots + \mathbf{P}_\ell.$$

Note that the algebraic multiplicity of λ_j as an eigenvalue of $\mathbf{T}(z)$ is the same as the algebraic multiplicity of λ_j as an eigenvalue of A. Recall that $m_1, \ldots, m_\ell$ are the algebraic multiplicities of $\lambda_1, \ldots, \lambda_\ell$, respectively. If λ_j is a semisimple eigenvalue of $\mathbf{T}(z)$ then λ_j is a semisimple eigenvalue of A. Hence there exist full column rank matrices $X_j \in \mathbb{C}^{r\times m_j}$ and $Y_j \in \mathbb{C}^{r\times m_j}$ such that $Y_j^\top X_j = I_{m_j}$, $\mathrm{span}(X_j) = N(A - \lambda_j I_{m_j})$, $\mathrm{span}(Y_j) = N((A - \lambda_j I_{m_j})^\top)$ and $P(\lambda_j, A) = X_j Y_j^\top$. This shows that $\mathbf{P}_j = CX_jY_j^\top B$. It follows that CX_j has full column rank and $Y_j^\top B$ has full row rank. Indeed, let $CX_ju = 0$. Now $AX_j = \lambda_j X_j \implies (A - \lambda_j I_{m_j})X_j u = 0$. Consequently, we have

$$\begin{bmatrix} A - \lambda_j I_{m_j} \\ C \end{bmatrix} X_j u = 0.$$

Since (C, A) is observable, we have $X_j u = 0 \Longrightarrow u = 0$ which shows that CX_j has full column rank. Similarly, controllability of (A, B) shows that $Y^\top B$ has full row rank. Hence by Theorem 5.2.10, we have

$$R(\mathbf{P}_j) = \text{span}(CX_j) = N(\mathbf{T}(\lambda_j)) \text{ and } R(\mathbf{P}_j^\top) = \text{span}(B^\top Y_j) = N(\mathbf{T}(\lambda_j)^\top). \quad (6.10)$$

This shows that $\mathbf{P}_j x$ (resp., $\mathbf{P}_j^\top y$) is a right (resp., left) eigenvectors of $\mathbf{T}(z)$ corresponding to λ_j whenever $\mathbf{P}_j x \neq 0$ (resp., $\mathbf{P}_j^\top y \neq 0$).

The next result extracts $\mathbf{P}_j$ from the matrix moments $S_0, \ldots, S_{\ell-1}$, which can be utilized to recover eigenvectors of $\mathbf{T}(z)$ from those of $\widehat{H}_\ell - \lambda H_\ell$.

Theorem 6.2.5. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Define*

$$\mathbf{V} := \begin{bmatrix} S_0 & \cdots & S_{\ell-1} \end{bmatrix} \in \mathbb{C}^{n\times n\ell} \text{ and } \mathbf{U} := \begin{bmatrix} S_0^\top & \cdots & S_{\ell-1}^\top \end{bmatrix} \in \mathbb{C}^{n\times n\ell}.$$

Suppose that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ are semisimple. Set $\alpha_j := q_{v_j}(\lambda_j)$, where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ for $j = 1 : \ell$. Then for $j = 1 : \ell$,

$$\mathbf{P}_j = \mathbf{V}(v_j \otimes I_n)/\alpha_j \text{ and } \mathbf{P}_j^\top = \mathbf{U}(v_j \otimes I_n)/\alpha_j$$

are residues of $\mathbf{T}(z)^{-1}$ and $(\mathbf{T}(z)^{-1})^\top$ at λ_j, respectively.

Define $\mathbf{v}_j := \mathbf{V}(v_j \otimes v) = \begin{bmatrix} S_0 v \cdots S_{\ell-1} v \end{bmatrix} v_j$ and $\mathbf{u}_j := \mathbf{U}(v_j \otimes u) = \begin{bmatrix} S_0^\top u \cdots S_{\ell-1}^\top u \end{bmatrix} v_j$ for $v, u \in \mathbb{C}^n$. Then $\mathbf{v}_j$ and $\mathbf{u}_j$, respectively, are right and left eigenvectors of $\mathbf{T}(z)$ corresponding to λ_j whenever $\mathbf{v}_j \neq 0$ and $\mathbf{u}_j \neq 0$.

Proof. Let $\mathbf{T}(z)^{-1} \simeq_{\mathcal{O}} C(zI_r - A)^{-1}B$ be a local minimal realization. Then by Theorem 5.2.10 $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalue of A. Let $P(\lambda_j, A)$ be the spectral projection of A coressponding to λ_j for $j = 1 : \ell$. Then $A = \lambda_1 P(\lambda_1, A) + \cdots + \lambda_\ell P(\lambda_\ell, A)$ and

$$(zI_r - A)^{-1} = \frac{P(\lambda_1, A)}{z - \lambda_1} + \cdots + \frac{P(\lambda_\ell, A)}{z - \lambda_\ell}. \quad (6.11)$$

By (6.9), we have $\mathbf{P}_j = CP(\lambda_j, A)B$ for $j = 1 : \ell$. Hence by (6.9) and (6.11), we have

$$S_p = \frac{1}{2\pi i}\int_\Gamma z^p \mathbf{T}(z)^{-1} dz = \sum_{j=1}^{\ell} \lambda_j^p \mathbf{P}_j = \begin{bmatrix} \mathbf{P}_1 & \cdots & \mathbf{P}_\ell \end{bmatrix} \begin{bmatrix} \lambda_1^p \\ \vdots \\ \lambda_\ell^p \end{bmatrix} \otimes I_n.$$

This shows that $\mathbf{V} = \begin{bmatrix} \mathbf{P}_1 & \cdots & \mathbf{P}_\ell \end{bmatrix} (V_\ell^\top \otimes I_n) \Longrightarrow \mathbf{P}_j = \mathbf{V}(V_\ell^{-\top} e_j \otimes I_n)$, where V_ℓ is the $\ell \times \ell$ Vandermonde matrix given in Theorem 6.2.3. Now by (6.5), we have

$(\Lambda_\ell - \lambda_j I_\ell)V_\ell^\top v_j = 0$ which shows that $V_\ell^\top v_j = \alpha_j e_j \Longrightarrow V_\ell^{-\top} e_j = v_j/\alpha_j$ for $j = 1 : \ell$. Hence we have $\mathbf{P}_j = \mathbf{V}(v_j \otimes I_n)/\alpha_j$ for $j = 1 : \ell$. The proof for $\mathbf{P}_j^\top$ is immediate.

By (6.10) we have $R(\mathbf{P}_j) = N(\mathbf{T}(\lambda_j))$ and $R(\mathbf{P}_j^\top) = N(\mathbf{T}(\lambda_j)^\top)$ for $j = 1 : \ell$. Hence $\mathbf{P}_j w$ is an eigenvector of $\mathbf{T}(z)$ corresponding to λ_j provided that $\mathbf{P}_j w \neq 0$. Therefore

$$\alpha_j \mathbf{P}_j w = \mathbf{V}(v_j \otimes w) = \begin{bmatrix} S_0 w \cdots S_\ell w \end{bmatrix} v_j = \mathbf{v}_j$$

is a right eigenvector of $\mathbf{T}(z)$ corresponding to λ_j. Similarly, if $\mathbf{P}_j^\top u \neq 0$ then

$$\alpha_j \mathbf{P}_j^\top u = \mathbf{U}(v_j \otimes u) = \begin{bmatrix} S_0^\top u \cdots S_\ell^\top u \end{bmatrix} v_j = \mathbf{u}_j$$

is a left eigenvector of $\mathbf{T}(z)$ corresponding to λ_j. □

Corollary 6.2.6. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $\mathbf{T}(z)$. Let $\mathbf{P}_j$ be the residue of $\mathbf{T}(z)^{-1}$ at λ_j and let v_j be given by $v_j := \begin{bmatrix} \beta_1^j & \cdots & \beta_\ell^j \end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then for $j = 1 : \ell$,*

$$\begin{aligned} \mathbf{P}_j &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)\mathbf{T}(z)^{-1} dz\right)/\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix} 1 & z & \cdots & z^{\ell-1} \end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial and $\alpha_j := q_{v_j}(\lambda_j)$.

Proof. Since $\mathbf{P}_j = \mathbf{V}(v_j \otimes I_n)/\alpha_j$ and

$$\begin{aligned} \mathbf{V}(v_j \otimes I_n) &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right) \\ &= \frac{1}{2\pi i}\int_\Gamma \sum_{k=1}^{\ell} \beta_k^j z^{k-1}\mathbf{T}(z)^{-1} dz \\ &= \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)\mathbf{T}(z)^{-1} dz, \end{aligned}$$

the desired result follows. □

Theorem 6.2.5 shows that exactly ℓ matrix moments $S_0, \ldots, S_{\ell-1}$ are needed for recovering the ℓ residues $\mathbf{P}_1, \ldots, \mathbf{P}_\ell$ of $\mathbf{T}(z)$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_\ell$, respectively. Also, for a random choice of u and v in $\mathbb{C}^n$, the vectors $\mathbf{V}(v_j \otimes v)$ and $\mathbf{U}(v_j \otimes u)$ are likely to be nonzero. Hence each such pair of vectors would yield a pair of right and left eigenvectors of $\mathbf{T}(z)$ corresponding to λ_j. On the other hand, if all linearly independent right and left eigenvectors of $\mathbf{T}(z)$ corresponding to λ_j are required then rank revealing QR factorization or SVD of $\mathbf{V}(v_j \otimes v)$ and $\mathbf{U}(v_j \otimes u)$ can be utilized to extract the eigenvectors.

Assume that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $\mathbf{T}(z)$. Then approximating the matrix moments by numerical quadrature

$$S_p \approx \sum_{j=1}^{N} w_j z_j^p \mathbf{T}(z_j)^{-1} \text{ and } \frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)\mathbf{T}(z)^{-1}dz \approx \sum_{j=1}^{N} w_j q_{v_j}(z_j)\mathbf{T}(z_j)^{-1}$$

we have the following method.

Algorithm for eigenvector

Input: Nodes z_j and weights w_j for j$= 1 : N$.

Eigenpair (λ_j, v_j) of $\hat{H}_\ell - \lambda H_\ell$ for $j = 1 : \ell$.

Output: Orthonormal bases of $N(\mathbf{T}(\lambda_j))$ and $N(\mathbf{T}(\lambda_j)^*)$ for $j = 1 : \ell$.

For j=1:ℓ

Compute $\mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z}_\mathtt{k}) = \begin{bmatrix} 1 & \mathtt{z}_\mathtt{k} & \cdots & \mathtt{z}_\mathtt{k}^{\ell-1} \end{bmatrix} \mathtt{v}_\mathtt{j}$ $\mathtt{for\ k = 1{:}N}$

Compute $\mathtt{A}_\mathtt{j} = \mathtt{sum}_{\mathtt{k}=1}^{\mathtt{N}} \mathtt{w}_\mathtt{k}\, \mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z}_\mathtt{k})\mathbf{T}(\mathtt{z}_\mathtt{k})^{-1}$

Compute compact SVD $[\mathtt{X}_\mathtt{j}, \mathtt{D}_\mathtt{j}, \mathtt{Y}_\mathtt{j}] = \mathtt{svd}(\mathtt{A}_\mathtt{j})$

Return the isometries $\mathtt{X}_\mathtt{j}$ and $\mathtt{Y}_\mathtt{j}$

End

The algorithm returns the isometries X_j and Y_j such that $R(X_j) = N(\mathbf{T}(\lambda_j))$ and $R(Y_j) = N(\mathbf{T}(\lambda_j)^*)$ for $j = 1 : \ell$. Thus columns of X_j and Y_j, respectively, are orthonormal right and left eigenvectors of $\mathbf{T}(z)$ corresponding to λ_j for $j = 1 : \ell$. We mention that the polynomial $\mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z})$ can be evaluated efficiently at $\mathtt{z}_\mathtt{k}$ by utilizing the MATLAB command `polyval`. Indeed, $\mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z}_\mathtt{k}) = \begin{bmatrix} 1 & \mathtt{z}_\mathtt{k} & \cdots & \mathtt{z}_\mathtt{k}^{\ell-1} \end{bmatrix} \mathtt{v}_\mathtt{j} = \mathtt{polyval}(\mathtt{flip}(\mathtt{v}_\mathtt{j}), \mathtt{z}_\mathtt{k})$. Also, $\mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z})$ can be evaluated at $\mathtt{z}_1, \ldots, \mathtt{z}_\mathtt{N}$ by a single call of `polyval` as follows. Set $\mathtt{zz} = [\mathtt{z}_1, \ldots, \mathtt{z}_\mathtt{N}]$ and compute $\mathtt{q}_\mathtt{j}\mathtt{zz} = \mathtt{polyval}(\mathtt{flip}(\mathtt{v}_\mathtt{j}), \mathtt{zz})$. Then $\mathtt{q}_\mathtt{j}\mathtt{zz}$ is a vector whose $\mathtt{k}$-th component is $\mathtt{q}_{\mathtt{v}_\mathtt{j}}(\mathtt{z}_\mathtt{k})$ for $\mathtt{k} = 1 : \mathtt{N}$.

6.3 Moment problems for operator-valued functions

Let X be a complex Banach space and $\mathbf{T} \in \mathbb{H}(\Omega, L(X))$ be regular. As in Chapter 3, let $\sigma(\mathbf{T})$ and $\sigma_d(\mathbf{T})$ denote the spectrum and discrete spectrum of $\mathbf{T}(z)$, respectively, and $\rho(\mathbf{T})$ be the resolvent set of $\mathbf{T}(z)$. Let $\mu \in \sigma_d(\mathbf{T})$. Then by Theorem 3.3.6,

$$\mathbf{T}(z) \sim_\mu D_\mu(z) := P_0 + (z-\mu)^{m_1} P_1 + \cdots + (z-\mu)^{m_r} P_r, \tag{6.12}$$

where P_j is a projection, $\mathrm{rank}(P_j) = 1$ for $j = 1 : r$ and $P_i P_j = 0$ for $i \neq j$. The positive integers $m_1 \leq \cdots \leq m_r$ are called the partial multiplicities of μ. The algebraic multiplicity of μ is defined as

$$m(\mu, T) := m_1 + \cdots + m_r.$$

If $\sigma \subset \sigma_d(\mathbf{T})$ is a finite set then $m(\sigma, \mathbf{T}) := \sum_{\mu\in\sigma} m(\mu, \mathbf{T})$ is the total algebraic multiplicity of the eigenvalues in σ.

Let $\mu \in \sigma_d(\mathbf{T})$. Then by Lemma 3.3.4, $\mathbf{T}(z)^{-1}$ is finitely meromorphic at $\mu \in \sigma_d(\mathbf{T})$ and $\int_{\Gamma_\mu} \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz$ is a finite rank operator, where Γ_μ is a simple closed curve in $\rho(\mathbf{T})$ isolating μ from the rest of $\sigma(\mathbf{T})$. Further, as $\mathbf{T}(z) \sim_\mu D_\mu(z)$, we have

$$\mathrm{Tr}\left(\int_{\Gamma_\mu} \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right) = \mathrm{Tr}\left(\int_{\Gamma_\mu} D_\mu(z)^{-1}D'_\mu(z)dz\right) \tag{6.13}$$

Hence by (6.12), we have $m(\mu, \mathbf{T}) = m_1 + \cdots + m_r = \dfrac{1}{2\pi i}\mathrm{Tr}\left(\int_{\Gamma_\mu} \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right)$.

Indeed, we have $D_\mu(z) = P_0 + (z-\mu)^{m_1}P_1 + \cdots + (z-\mu)^{m_r}P_r$ and hence

$$D_\mu(z)^{-1}D'_\mu(z) = m_1(z-\lambda_1)^{-1}P_1 + \cdots + m_\ell(z-\lambda_\ell)^{-1}P_\ell.$$

Consequently, for $g \in \mathbb{H}(\Omega)$, we have

$$\mathrm{Tr}\left(\int_{\Gamma_\mu} g(z)\mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right) = \mathrm{Tr}\left(\int_{\Gamma_\mu} g(z)D_\mu(z)^{-1}D'_\mu(z)dz\right) = m(\mu, \mathbf{T})g(\mu) \tag{6.14}$$

Hence for $g(z) = 1$ we have the desired result.

Let $\Gamma \subset \rho(\mathbf{T})$ be a positively oriented rectifiable simple closed curve. Set $\mathcal{O} := \mathrm{Int}(\Gamma)$ and assume that $\mathcal{O} \subset \Omega$. Suppose that $\sigma(\mathbf{T}) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(\mathbf{T})$. Then $\mathbf{T}(z)^{-1}$ is finitely meromorphic at λ_j for $j = 1 : \ell$. Let m_j be the algebraic multiplicity of λ_j for $j = 1 : \ell$. Then we have the following result which is an analogue of Theorem 6.2.1.

Theorem 6.3.1. *Let* $\mathrm{mean}(\lambda_1, \ldots, \lambda_\ell) := \dfrac{m_1\lambda_1 + \cdots + m_\ell\lambda_\ell}{m_1 + \cdots + m_\ell}$ *be the arithmetic mean of the eigenvalues* $\lambda_1, \ldots, \lambda_\ell$ *of* $\mathbf{T}(z)$*. Let* $g \in \mathbb{H}(\Omega)$*. Then we have*

$$\mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma g(z)\mathbf{T}^{-1}(z)\mathbf{T}'(z)dz\right) = m_1 g(\lambda_1) + \cdots + m_\ell g(\lambda_\ell). \tag{6.15}$$

In particular, we have

$$\mathrm{mean}(\lambda_1, \ldots, \lambda_\ell) = \frac{\mathrm{Tr}\left(\int_\Gamma z\, \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right)}{\mathrm{Tr}\left(\int_\Gamma \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right)}.$$

Proof. Let $\Gamma_j \subset \rho(\mathbf{T})$ be a simple closed curve isolating λ_j from the rest of $\sigma(\mathbf{T})$. Note that $\mathbf{T}(z) \sim_{\lambda_j} D_j(z)$, where $D_j(z)$ is of the form (6.12). Then by (6.14), we have

$$\mathrm{Tr}\left(\frac{1}{2\pi i}\int_{\Gamma_j} g(z)\mathbf{T}^{-1}(z)\mathbf{T}'(z)\right) = m_j g(\lambda_j) \text{ for } j = 1 : \ell.$$

Consequently, we have

$$\begin{aligned}\mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma g(z)\mathbf{T}^{-1}(z)\mathbf{T}'(z)dz\right) &= \sum_{j=1}^{\ell}\mathrm{Tr}\left(\frac{1}{2\pi i}\int_{\Gamma_j} g(z)\mathbf{T}^{-1}(z)\mathbf{T}'(z)dz\right)\\ &= m_1 g(\lambda_1) + \cdots + m_\ell g(\lambda_\ell).\end{aligned}$$

Finally, considering $g(z) = 1$ and $g(z) = z$ we obtain $\mathrm{mean}(\lambda_1, \ldots, \lambda_\ell)$. □

By Theorem 6.3.1, the p-th order tracial moment s_p is given by

$$s_p = \mathrm{Tr}\left(\frac{1}{2\pi i}\int_\Gamma z^p\, \mathbf{T}(z)^{-1}\mathbf{T}'(z)dz\right) = m_1\lambda_1^p + \cdots + m_\ell\lambda_\ell^p \text{ for } p \in \mathbb{Z}_+. \tag{6.16}$$

As before, construct $k \times k$ Hankel matrix H_k and shifted Hankel matrix $\widehat{H}_k$ from the tracial moments s_p for $p \in \mathbb{Z}_+$. Then by (6.16) we have the following result which is an analogue of Theoem 6.2.3 and whose proof is the same as that of Theorem 6.2.3.

Theorem 6.3.2. *Let* $\sigma(\mathbf{T}) \cap \mathrm{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(\mathbf{T})$ *and* $m_1, \ldots, m_\ell$ *be algebraic multiplicities of* $\lambda_1, \ldots, \lambda_\ell$*, respectively. Then* $\lambda_1, ..., \lambda_\ell$ *are the simple eigenvalues of the* $\ell \times \ell$ *Hankel pencil* $\widehat{H}_\ell - \lambda H_\ell$*. The algebraic multiplicities of the eigenvalues are given by the* $\ell \times \ell$ *Vandermonde system*

$$V_\ell \begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix}\begin{bmatrix} m_1 \\ \vdots \\ m_\ell \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ s_{\ell-1} \end{bmatrix}.$$

The number ℓ satisfies $\ell = \operatorname{rank}(H_\ell) = \operatorname{rank}(H_N) = \operatorname{rank}(H_{\ell+p})$ for any non-negative integer p, where $N := s_0 = m_1 + \cdots + m_\ell$.

Note that $\lambda_1, \ldots, \lambda_\ell$ are the eigenvalues of the matrix $A_\ell := H_\ell^{-1}\widehat{H}_\ell$. Now we define formal orthogonal polynomials through which we construct a monic polynomial which is the characteristic polynomial of A_ℓ.

Set $\sigma := \sigma(\mathbf{T}) \cap \operatorname{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(\mathbf{T})$. Consider the symmetric bilinear form $\langle \cdot, \cdot \rangle : \mathbb{C}[z] \times \mathbb{C}[z] \to \mathbb{C}$ given by

$$\big\langle \phi(z), \psi(z) \big\rangle := \operatorname{Tr}\left(\frac{1}{2\pi i}\int_\Gamma \phi(w)\psi(w)\mathbf{T}(w)^{-1}\mathbf{T}'(w)dw\right) \text{ for } \phi(z), \psi(z) \in \mathbb{C}[z].$$

Then observe that $m := \langle 1, 1 \rangle = m_1 + \cdots + m_\ell$ is the total algebraic multiplicity of the eigenvalues in σ and $\operatorname{mean}(\lambda_1, \ldots, \lambda_\ell) = \dfrac{\langle 1, z \rangle}{\langle 1, 1 \rangle}$.

Definition 6.3.3. *A monic polynomial $\phi_p(z)$ of degree p is said to be a formal orthogonal polynomial (FOP) of $\mathbf{T}(z)$ corresponding to σ if*

$$\langle z^k, \phi_p(z) \rangle = 0 \;\; for \;\; k = 0, 1, ..., p-1.$$

An FOP $\phi_p(z)$ is called regular if it is unique.

Theorem 6.3.4. *Let $\sigma := \sigma(\mathbf{T}) \cap \operatorname{Int}(\Gamma) = \{\lambda_1, \ldots, \lambda_\ell\} \subset \sigma_d(\mathbf{T})$ and $m_1, \ldots, m_\ell$ be algebraic multiplicities of $\lambda_1, \ldots, \lambda_\ell$, respectively. Consider the tracial moments s_p for $p \in \mathbb{Z}_+$. Then an FOP $\phi_p(z)$ exists uniquely $\iff$ H_p is invertible.*

In particular, the FOP

$$\phi_\ell(z) := a_0 + a_1 z + \cdots + a_{\ell-1}z^{\ell-1} + z^\ell$$

exists uniquely and is given by

$$H_\ell \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{\ell-1} \end{bmatrix} = \begin{bmatrix} s_0 & s_1 & \ldots & s_{\ell-1} \\ s_1 & s_2 & \ldots & s_\ell \\ \vdots & \vdots & \ldots & \vdots \\ s_{\ell-1} & s_\ell & \ldots & s_{2\ell-2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{\ell-1} \end{bmatrix} = - \begin{bmatrix} s_\ell \\ s_{\ell+1} \\ \vdots \\ s_{2\ell-1} \end{bmatrix}.$$

Further, $\lambda_1, \ldots, \lambda_\ell$ are simple roots of $\phi_\ell(z)$. In fact, $\phi_\ell(z) = \det(zI_\ell - H_\ell^{-1}\widehat{H}_\ell)$.

Proof. Consider $\phi_p(z) := a_0 + a_1 z + \cdots + a_{\ell-1}z^{p-1} + z^p$. Note that $\langle z^i, z^j\rangle = s_{i+j}$. Hence $\langle z^k, \phi_p(z)\rangle = \sum_{i=0}^{p-1} a_i \langle z^k, z^i\rangle + \langle z^k, z^p\rangle = \sum_{i=0}^{p-1} a_i s_{k+i} + s_{k+p}$ for $k = 0, ..., p-1$. Thus setting $\langle z^k, \phi_p(z)\rangle = 0$ for $k = 0, 1, \ldots, p-1$, we obtain the linear system

$$H_p \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix} = \begin{bmatrix} s_0 & s_1 & \ldots & s_{p-1} \\ s_1 & s_2 & \ldots & s_p \\ \vdots & \vdots & \ldots & \vdots \\ s_{p-1} & s_p & \ldots & s_{2p-2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix} = - \begin{bmatrix} s_p \\ s_{p+1} \\ \vdots \\ s_{2p-1} \end{bmatrix}. \tag{6.17}$$

The system (6.17) has a unique solution $\iff H_p$ is nonsingular. In particular, we know that H_ℓ is invertible. Hence $\phi_\ell(z)$ exists uniquely, that is, $\phi_\ell(z)$ is a regular FOP.

Next, by (6.15), we have $\langle \phi(z), \psi(z)\rangle = m_1\phi(\lambda_1)\psi(\lambda_1) + \cdots + \lambda_\ell\phi(\lambda_\ell)\psi(\lambda_\ell)$ for all $\phi(z), \psi(z) \in \mathbb{C}[z]$. Hence $\langle z^k, \phi_\ell(z)\rangle = m_1\lambda_1^k\phi_\ell(\lambda_1) + \cdots + m_\ell\lambda_\ell^k\phi_\ell(\lambda_\ell)$ for $k = 0 : \ell - 1$. Since $\langle z^k, \phi_\ell(z)\rangle = 0$ for $k = 0 : \ell - 1$, we have

$$V_\ell \begin{bmatrix} m_1\phi_\ell(\lambda_1) \\ \vdots \\ m_\ell\phi_\ell(\lambda_\ell) \end{bmatrix} = \begin{bmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_\ell \\ \vdots & \ldots & \vdots \\ \lambda_1^{\ell-1} & \ldots & \lambda_\ell^{\ell-1} \end{bmatrix} \begin{bmatrix} m_1\phi_\ell(\lambda_1) \\ \vdots \\ m_\ell\phi_\ell(\lambda_\ell) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since V_ℓ is nonsingular, we have $m_k\phi_\ell(\lambda_k) = 0 \Longrightarrow \phi_\ell(\lambda_k) = 0$ for $k = 0 : \ell - 1$. This shows that $\lambda_1, \ldots, \lambda_\ell$ are simple roots of $\phi_\ell(z)$.

Consider the companion matrix C_ℓ of $\phi_\ell(z)$ given by

$$C_\ell := \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & \ldots & \ldots & 1 & -a_{\ell-1} \end{bmatrix}.$$

It is well known that $\phi_\ell(z) = \det(zI_\ell - C_\ell)$. It is easy to see that $\widehat{H}_\ell = H_\ell C_\ell$ which shows that $C_\ell = H_\ell^{-1}\widehat{H}_\ell$ which in turn shows that $\phi_\ell(z) = \det(zI_\ell - H_\ell^{-1}\widehat{H}_\ell)$. □

As in the case of holomorphic matrices, eigenvectors of $\mathbf{T}(z)$ can be recovered from the operator moments. Indeed, define $\mathbf{M}(z) := \mathbf{T}(z)^{-1}$. Then $\mathbf{M}(z)$ is finitely meromorphic in $\mathcal{O}$. Let $\mathbf{P}_j := \mathrm{Res}(\lambda_j, \mathbf{M}(z))$ be the residue of $\mathbf{M}(z)$ at λ_j for $j = 1 : \ell$. Then it can be shown that

$$\mathbf{P}_j = \frac{1}{2\pi i}\int_{\Gamma_j} \mathbf{M}(z)dz \text{ and } R(\mathbf{P}_j) = N(\mathbf{T}(\lambda_j)),$$

where Γ_j is a curve in $\mathcal{O} \cap \rho(\mathbf{T})$ isolating λ_j from the rest of $\sigma(\mathbf{T})$ for $j = 1 : \ell$. Further, it can be shown that for $p \in \mathbb{Z}_+$,

$$S_p = \frac{1}{2\pi i}\int_{\Gamma} z^p \mathbf{M}(z)dz = \sum_{j=1}^{\ell} \frac{1}{2\pi i}\int_{\Gamma_j} z^p \mathbf{M}(z)dz = \lambda_1^p \mathbf{P}_1 + \cdots + \lambda_\ell^p \mathbf{P}_\ell.$$

Consequently, we have the following result which extracts $\mathbf{P}_j$ from the operator moments $S_0, \ldots, S_{\ell-1}$ and can be utilized to recover eigenvectors of $\mathbf{T}(z)$ from those of $\widehat{H}_\ell - \lambda H_\ell$.

Theorem 6.3.5. *Let $(\lambda_1, v_1), \ldots, (\lambda_\ell, v_\ell)$ be eigenpairs of $\widehat{H}_\ell - \lambda H_\ell$. Suppose that $\lambda_1, \ldots, \lambda_\ell$ are semisimple eigenvalues of $\mathbf{T}(z)$. Let $\mathbf{P}_j$ be the residue of $\mathbf{T}(z)^{-1}$ at λ_j and let v_j be given by $v_j := \begin{bmatrix}\beta_1^j & \cdots & \beta_\ell^j\end{bmatrix}^\top \in \mathbb{C}^\ell$ for $j = 1 : \ell$. Then for $j = 1 : \ell$,*

$$\begin{aligned} \mathbf{P}_j &= \left(\beta_1^j S_0 + \cdots + \beta_\ell^j S_{\ell-1}\right)/\alpha_j \\ &= \left(\frac{1}{2\pi i}\int_\Gamma q_{v_j}(z)\mathbf{T}(z)^{-1}dz\right)/\alpha_j, \end{aligned}$$

where $q_{v_j}(z) := \begin{bmatrix}1 & z & \cdots & z^{\ell-1}\end{bmatrix} v_j \in \mathbb{C}[z]$ is a polynomial and $\alpha_j := q_{v_j}(\lambda_j)$.

For $j = 1 : \ell$, define $\mathbf{v}_j := \begin{bmatrix}S_0 v \cdots S_{\ell-1} v\end{bmatrix} v_j$ for any $v \in \mathbb{C}^n$. Then $\mathbf{v}_j$ is an eigenvector of $\mathbf{T}(z)$ corresponding to λ_j whenever $\mathbf{v}_j \neq 0$.

6.4 Computation

Let $\mathbf{T} \in \mathbb{H}(\Omega)^{n\times n}$ be regular. We have seen that the eigenvalues $\lambda_1, \ldots, \lambda_\ell$ of $\mathbf{T}(z)$ inside the region $\mathcal{O} := \text{Int}(\Gamma)$ can be computed from the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ constructed from the tracial moments s_p of $\mathbf{T}(z)$. The corresponding eigenvectors can be recovered from the eigenvectors of $\widehat{H}_\ell - \lambda H_\ell$ by utilizing matrix moments S_p for $p \in \mathbb{Z}_+$. Thus approximating the moments s_p and S_p by numerical quadrature

$$\begin{aligned} s_p &\approx \sum_{j=1}^{N} z_j^p w_j \text{Tr}(\mathbf{T}(z_j)^{-1}\mathbf{T}'(z_j)) \\ S_p &\approx \sum_{j=1}^{N} z_j^p w_j \mathbf{T}(z_j)^{-1} \end{aligned}$$

with nodes $z_1, \ldots, z_N$ and weights $w_1, \ldots, w_N$, we obtain a numerical algorithm. We illustrate performance of the resulting algorithm by considering a few examples.

Hankel eigenvalue problems such as $\ell \times \ell$ eigenvalue problem $(\widehat{H}_\ell - \lambda H_\ell)v = 0$ are known to be potentially ill-conditioned when ℓ is large. It is therefore advisable to compute a small number of eigenvalues of $\mathbf{T}(z)$ at a time. On the other hand, when ℓ is large, the computed eigenvalues can be refined by Newton iteration to improve the accuracy.

6.4.1 Newton iteration

Let μ be a computed eigenvalue of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$. Let $f(z) := \det(\mathbf{T}(z))$. Then μ is an approximate zero $f(z)$, that is, $f(\mu) \approx 0$. So, the Newton's method can be used to refine μ. We have seen that by the Jacobi formula

$$\frac{f'(z)}{f(z)} = \mathrm{Tr}(\mathbf{T}^{-1}(z)\mathbf{T}'(z)).$$

Setting $\lambda_0 := \mu$, the Newton's method is given by

$$\lambda_{j+1} = \lambda_j - \frac{f(\lambda_j)}{f'(\lambda_{j+1})} = \lambda_j - \frac{1}{\mathrm{Tr}(\mathbf{T}^{-1}(\lambda_j)\mathbf{T}'(\lambda_j))}.$$

Our numerical method also provides the multiplicity of μ. So, when μ is a multiple eigenvalue of $\mathbf{T}(z)$ with multiplicity m, we consider the Newton iteration

$$\lambda_{j+1} = \lambda_j - \frac{m}{\mathrm{Tr}(\mathbf{T}^{-1}(\lambda_j)\mathbf{T}'(\lambda_j))} \tag{6.18}$$

to refine the eigenvalue μ.

6.4.2 Backward error

For accuracy assessment of an approximate eigenvalue μ of $\mathbf{T}(z)$ we compute the backward error of μ. Assume that $\mathbf{T}(z)$ is given by

$$\mathbf{T}(z) := f_1(z)A_1 + \cdots + f_m(z)A_m$$

for some functions $f_1, \ldots, f_m$ in $\mathbb{H}(\Omega)$ and fixed matrices $A_1, \ldots, A_m$ in $\mathbb{C}^{n\times n}$. Now, consider perturbation $\Delta\mathbf{T}(z)$ of the form

$$\Delta\mathbf{T}(z) := f_1(z)\Delta A_1 + \cdots + f_m(z)\Delta A_m,$$

where $\Delta A_1, \ldots, \Delta A_m$ are matrices in $\mathbb{C}^{n\times n}$. Then the backward error μ is defined as

$$\eta(\mu, \mathbf{T}) := \min\left\{\sqrt{\|\Delta A_1\|_2^2 + \cdots + \|\Delta A_m\|_2^2} : \det(\mathbf{T}(\mu) + \Delta\mathbf{T}(\mu)) = 0\right\}.$$

It can be shown that (see, [22])

$$\eta(\mu, \mathbf{T}) = \frac{\sigma_{\min}(\mathbf{T}(\mu))}{\|(f_1(\mu), \ldots, f_m(\mu))\|_2},$$

where $\sigma_{\min}(\mathbf{T}(\mu))$ is the smallest singular value of the matrix $\mathbf{T}(\mu)$. If $\eta(\mu, \mathbf{T})$ is small then μ is accepted as an eigenvalue of $\mathbf{T}(z)$.

6.4.3 Numerical Examples

We now consider a few numerical examples to illustrate the performance our method. We approximate contour integrals by trapezoid method. All the computations are performed in MATLAB and the examples are taken from the nonlinear eigenvalue problem MATLAB toolbox [10]

Example 6.4.1 (qep,[10]). *Consider $Q(\lambda) := \lambda^2 M + \lambda C + K$, where*

$$M = \begin{bmatrix} 0 & 6 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & -6 & 0 \\ 2 & -7 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The six eigenvalues of $Q(\lambda)$ are $\infty, -i, i, 1, \frac{1}{3}$ and $\frac{1}{2}$. We consider Γ to be the circle of radius 1.5 and centered at the origin $0 + 0i$. Table 6.1 gives eigenvalues of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ for $N = 100$.

Computed Eigenvalues	Absolute Errors	Backward Errors
1.57378568227378e-15-1.00000000000001e+00i	6.41334434827567e-15	7.40549217176541e-15
-6.88373327490615e-15+1.00000000000000e+00i	7.38143156488219e-15	8.52334300197907e-15
1.00000000000006e+00+3.91209626500391e-14i	6.95542836036101e-14	2.84704145948781e-14
3.33333333336468e-01+2.21436492351705e-12i	3.83808447343764e-12	1.14503837569542e-12
5.00000000003777e-01+2.57760136781624e-12i	4.57233449156501e-12	7.82612809704569e-13

Table 6.1: Center $c = 0 + i0$, radius $r = 1.5$ and # nodes $N = 100$

Example 6.4.2 (Butterfly, [10]). *This is a quartic T-even matrix polynomial*

$$P(\lambda) := \lambda^4 A_4 + \lambda^2 A_3 + \lambda^2 A_2 + \lambda A_1 + A_0$$

of dimension m^2 and depends on a 10×1 parameter vector c. Its spectrum has a butterfly shape. The coefficient matrices are Kronecker products with A_4 and A_2 real and symmetric and A_3 and A_1 real skew-symmetric, assuming c is real. The default is $m = 8$.The coefficient matrices are generated from MATLAB *NLEVP toolbox [10].*

We consider Γ to be the circle of radius $r = 0.22$ and centered at $c = -1 - 0.5i$. Table 6.2 gives the eigenvalues of of the Hankel pencil $\widehat{H}_\ell - \lambda H_\ell$ for $N = 200$. We find that all computed eigenvalues are simple.

Problem: Butterfly

Computed Eigenvalues	Backward Errors
-1.02618997320756e+00-6.85703044215422e-01i	1.71914597380728e-12
-8.64617980463986e-01-6.51815654485066e-01i	3.82580443847945e-11
-9.53854042153464e-01-6.11439886639829e-01i	5.51162838753142e-09
-9.94127905252742e-01-5.35135844241971e-01i	9.16333532646743e-08
-8.53220324727572e-01-5.17099952102852e-01i	5.44292866991764e-08
-9.30913090447226e-01-4.80355996605403e-01i	6.91656032720915e-06
-8.36379010692839e-01-4.04087636721792e-01i	4.15642326925671e-08
-9.67237603495008e-01-4.30157331331871e-01i	5.38944138591991e-06
-9.10092484618337e-01-3.84283030851018e-01i	1.27908191579372e-05
-9.47628789624199e-01-3.60769566015209e-01i	4.98104057130454e-06
-8.94552027538694e-01-3.19324913427143e-01i	4.38437695933354e-08
-9.35858730014987e-01-3.21436784098499e-01i	2.57333233818091e-07

Table 6.2: Center $c = -1 - 0.5i$, radius $r = 0.22$, and # nodes $N = 200$

Observe that the computed eigenvalues of the butterfly problem have backward errors of order 10^{-5} or less. We perform Newton iteration on these eigenvalues to improve the accuracy. Table 6.3 shows the results after one step of Newton iteration and Table 6.4 shows the results after two steps of Newton iteration. As expected the accuracy increases significantly.

Problem: Butterfly

Iterated Eigenvalues (Iteration 1)	Backward Error
-1.02618997320821e+00-6.85703044215535e-01i	8.55883097528700e-17
-8.64617980453660e-01-6.51815654480523e-01i	3.99530167760573e-17
-9.53854040217057e-01-6.11439886362263e-01i	1.84418256477503e-15
-9.94127888031000e-01-5.35135868221294e-01i	4.97641072551466e-13
-8.53220341698702e-01-5.17099955869029e-01i	1.08077102145015e-13
-9.30912753585458e-01-4.80358607987226e-01i	2.56445338606946e-09
-8.36379010253743e-01-4.04087655723411e-01i	7.00566088005300e-14
-9.67236650972395e-01-4.30155816400882e-01i	1.59296907884947e-09
-9.10093928782547e-01-3.84276901205813e-01i	9.33721084873891e-09
-9.47627895889479e-01-3.60767903899836e-01i	1.47122453671497e-09
-8.94552041898777e-01-3.19324937764473e-01i	1.23848078031660e-13
-9.35858719011995e-01-3.21436873529427e-01i	3.31632300357617e-12

Table 6.3: Iterated eigenvalues and backward errors (Iteration 1)

The backward errors of the first three eigenvalues in Table 6.3 is of order $\mathcal{O}(10^{-16})$ after first iteration. So we need not refine these eigenvalues. Table 6.4 shows the results after two steps of Newton iterations.

Problem: Butterfly

Iterated Eigenvalues (Iteration 2)	Backward Error
-1.02618997320821e+00-6.85703044215535e-01i	1.19595687952740e-16
-8.64617980453660e-01-6.51815654480523e-01i	5.58278173843677e-17
-9.53854040217056e-01-6.11439886362263e-01i	3.17807975881956e-16
-9.94127888031145e-01-5.35135868221436e-01i	1.04548405409803e-16
-8.53220341698647e-01-5.17099955869046e-01i	3.22158406907252e-16
-9.30912754988716e-01-4.80358607551854e-01i	8.89533215262868e-16
-8.36379010253795e-01-4.04087655723379e-01i	3.53532429636910e-17
-9.67236651050897e-01-4.30155815626538e-01i	3.72751991139599e-16
-9.10093936597451e-01-3.84276901220163e-01i	2.01142944982932e-14
-9.47627896206371e-01-3.60767903058135e-01i	4.74106400399762e-16
-8.94552041898822e-01-3.19324937764334e-01i	1.80691353971286e-16
-9.35858719013910e-01-3.21436873529878e-01i	9.86602811450068e-17

Table 6.4: Iterated eigenvalues and backward errors (Iteration 2)

We now consider holomorphic eigenvalue problems that arise from delay-differential equations of the form $\frac{dx}{dt} = Bx(t) + Ax(t - \tau)$, where $\tau > 0$ is a time-delay parameter.

Example 6.4.3 (Time_delay, [10]). *Consider* $\mathbf{T}(\lambda) := -\lambda I_3 + A_0 + A_1 e^{-\lambda}$, *where*

$$A_0 := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{bmatrix} \text{ and } A_1 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -b_3 & -b_2 & -b_1 \end{bmatrix}.$$

Here $a_1 \approx 3.98, a_2 \approx 108, a_3 \approx 531,$ *and* $b_1 \approx 13.6, b_2 \approx 18.7, b_3 \approx 1363$. *The coefficient matrices are generated using* MATLAB *NLEVP toolbox [10].*

This problem has a double defective eigenvalue at $\lambda := 3\pi i$. The computed results show 8 distinct eigenvalues. Among these, the computed eigenvalues

$$-8.769e\text{-}11 + 9.4248 \approx 3\pi i \text{ and } 1.2858e\text{-}11 - 9.4248i \approx -3\pi i$$

are the results of splitting of λ due to approximation. Even without Newton iteration, the backward errors of these eigenvalues are very small.

We consider Γ to be the circle centered at $c = 0 + i0$ and radius $r = 22.0$. For $N = 200$, the computed eigenvalues are given in Table 6.5.

Problem: Time_delay

Computed Eigenvalues	Backward Errors
-4.22996397305367e-01 + 2.04853626079604e+01i	2.80233085924428e-15
-3.96573424780551e-11 + 1.41371669411581e+01i	1.71512374631199e-13
-1.15270625550560e-10 + 9.42477796076992e+00i	1.11147991853796e-16
7.05244108642892e-01 + 2.74146676197162e+00i	1.45388622820372e-10
7.05244108858527e-01 - 2.74146676258927e+00i	1.27925250565019e-10
-1.55698546676974e-11 - 9.42477796083883e+00i	1.33544954514744e-16
-6.25551389012847e-12 - 1.41371669411742e+01i	9.06593511255611e-14
-4.22996397305057e-01 - 2.04853626079605e+01i	1.61043003429410e-15

Table 6.5: Center $c = 0 + 0i$, Radius $r = 22.00$ and # Nodes $N = 200$

The backward errors of 4th and 5th eigenvalues in Table 6.5 is of order $\mathcal{O}(10^{-10})$. We refine these two eigenvalues using Newton iteration for reducing the backward errors. Table 6.6 gives the results after one step of Newton iteration.

Iterated Eigenvalues	Backward Errors
7.05244109106679e-01 + 2.74146676220549e+00i	3.42552350761054e-16
7.05244109106679e-01 - 2.74146676220549e+00i	2.77660666816832e-16

Table 6.6: 4th and 5th eigenvalues and backward errors (Iteration 1)

Example 6.4.4. *[10, Time_delay2]. Consider* $F(\lambda) := \lambda I_2 + B_0 + A_1 e^{-\tau\lambda}$ *with*

$$B_0 = \begin{bmatrix} 5 & -1 \\ -2 & 6 \end{bmatrix} \quad and \quad A_1 = \begin{bmatrix} 2 & -1 \\ -4 & 1 \end{bmatrix}.$$

The default value of τ *is* 1.

We consider Γ to be the circle with center $c = 0 + i0$ and radius $r = 10$. For $N = 200$, the computed results shows that there are 7 distinct eigenvalues inside the circle. Table 6.7 shows the computed eigenvalues.

Problem: Time_delay2

Computed Eigenvalues	Backward Errors
-1.05804451367999e+00 - 8.44995491281638e+00i	7.07183038027321e-11
-2.26740253701666e+00 - 5.06926669987741e+00i	1.03195870131910e-09
-6.35474596890421e-01 - 2.71752199606495e+00i	1.37340931210166e-08
-1.53587609008173e+00 + 5.09488767432416e-11i	1.10446099330149e-08
-6.35474596884534e-01 + 2.71752199610630e+00i	1.37783491039590e-08
-2.26740253701299e+00 + 5.06926669987787e+00i	1.03296718415130e-09
-1.05804451368002e+00 + 8.44995491281636e+00i	7.07289473580911e-11

Table 6.7: center $c = 0 + 0i$, radius $= 10$, # nodes $N = 200$.

We perform Newton iteration to reduce the backward errors. Table 6.8 shows the results after one step of Newton iteration.

Iterated Eigenvalues	Backward Errors
-1.05804451362771e+00 - 8.44995491276330e+00i	8.14653408536819e-16
-2.26740253833744e+00 - 5.06926669783878e+00i	9.53739580044737e-17
-6.35474591311729e-01 - 2.71752198972701e+00i	8.85554287569495e-17
-1.53587607147439e+00 + 2.03903932007655e-18i	2.60905556137891e-16
-6.35474591311729e-01 + 2.71752198972701e+00i	8.85554287569495e-17
-2.26740253833744e+00 + 5.06926669783878e+00i	9.53739580044737e-17
-1.05804451362771e+00 + 8.44995491276330e+00i	8.14653408536819e-16

Table 6.8: Eigenvalues and backward errors (Iteration 1)

The numerical examples considered above illustrate that our method can be gainfully utilized to compute a small number of eigenvalues of $\mathbf{T}(z)$. For the numerical quadrature, it is enough to consider $N = 100$ or $N = 200$ nodes. If higher accuracy is desired then a few steps of Newton iteration can be used to achieve the desired accuracy. Thus our method together with Newton iteration can be used to solve the NEP $\mathbf{T}(\lambda)v = 0$. The main advantage of our method is that the size of the Hankel eigenvalue problem to be solved is determined by the number of distinct eigenvalues ℓ inside the curve Γ irrespective of the total algebraic multiplicity of the eigenvalues. So, if ℓ is small then we have to solve only a small $\ell \times \ell$ eigenvalue problem $(\widehat{H}_\ell - \lambda H_\ell)v = 0$.

7

Conclusions

The main aim of the thesis was to study theory and numerics of holomorphic and meromorphic eigenvalue problems. We have considered several broad problems, namely, spectral analysis of holomorphic and meromorphic matrices, spectral perturbation theory for holomorphic matrices, realization and linearization of meromorphic matrices, and spectral recovery methods for solution of nonlinear eigenvalue problems.

Canonical forms are powerful tools for spectral analysis. Smith forms of matrix polynomials and Smith-McMillan forms of rational matrices are well known canonical forms which are used extensively in theory and applications. Although, local canonical forms such as local Smith forms of holomorphic/meromorphic matrices are well known in the literature, global canonical forms such as Smith forms/Smith-McMillan forms of holomorphic/meromorphic matrices are not available in the literature. We have filled this gap. We have proved, among other things, global Smith forms/Smith-McMillan forms of holomorphic/meromorphic matrices which are akin to Smith forms/Smith-McMillan forms of matrix polynomials/rational matrices.

Linear systems theoretic analysis (e.g., MFDs, realization theory, and system matrices) of rational matrices are well known in the literature in which Smith-McMillan form of rational matrices plays an important role. We have comprehensively developed linear system theoretic analysis of meromorphic matrices on the same lines as that of rational matrices. We have utilized global Smith-McMillan form to develop global MFDs, realization theory, and analytic system matrices of meromorphic matrices. We have established relationship between canonical forms of meromorphic matrices with those of MFDs, state matrices, and analytic system matrices. These results are akin to corresponding results for rational matrices.

Spectral perturbation theory for nonlinear eigenvalue problem is a challenging task.

We have studied spectral perturbation theory for holomorphic operator-valued functions by utilizing local linearizations of holomorphic operator-valued functions. We have derived, among other things, perturbation bounds for discrete eigenvalues (and their corresponding eigenvectors) of holomorphic operator-valued functions.

Realization theory of rational matrices is a classical topic which is used extensively in linear systems theory. We have developed a local realization theory, which is akin to the realization theory of rational matrices, for meromorphic matrices. We have shown that a minimal local realization of a meromorphic matrix can be computed by utilizing Markov parameters. Consequently, poles of a meromorphic matrix can be computed from a minimal local realization. Further, we have introduced the concept of local linearization of a meromorphic matrix and, by utilizing a minimal local realization, we have constructed a local linearization which subsumes the local linearization of a holomorphic matrix. Thus, spectral perturbation bounds for meromorphic matrices can be easily derived from their holomorphic counterparts by considering a local linearization.

Solving a nonlinear eigenvalue problem (NEP) numerically is a challenging task. We have shown that solving an NEP is equivalent to determining poles of an appropriate meromorphic matrix and that poles of a meromorphic matrix can be computed from a minimal local realization of the meromorphic matrix. Thus, we have shown that an NEP can be solved by a minimal local realization. Also, we have proposed spectral recovery problems for linear and nonlinear eigenvalue problems and provided their solutions. We have shown that solutions to spectral recovery problems provide alternative numerical methods for solving an NEP in which the size of the eigenvalue problems to be solved is the same as the number of distinct eigenvalues sought to be determined irrespective of the algebraic multiplicities of the eigenvalues.

Finally, we mention that numerical issues such as the conditioning of the problems considered and the stability analysis of the numerical methods proposed in the thesis require a detailed critical analysis. Moreover, a robust implementation of the numerical methods proposed in the thesis and their performance analysis also require further investigation. We plan to investigate all these issues in a future work.

Publication from the thesis

Published paper:

- Rafikul Alam and Jibrail Ali, *Perturbations of discrete spectra of holomorphic operator-valued functions*, The Journal of Analysis, 29(2021), pp.551-569.
- Rafikul Alam and Jibrail Ali, *Canonical forms of holomorphic and meromorphic matrices,* Contemp. Math., AMS, to appear.

Preprints/papers under preparation:

1. Rafikul Alam and Jibrail Ali, *System-theoretic analysis of meromorphic matrices,* preprint.
2. Rafikul Alam and Jibrail Ali, *Spectral recovery problems for linear operators,* preprint.
3. Rafikul Alam and Jibrail Ali, *Realization and linearization of meromorphic matrices with application to NEP,* in preparation.
4. Rafikul Alam and Jibrail Ali, *Moment problems for eigenvalues and eigenspaces of holomorphic matrices,* in preparation.

Bibliography

[1] R. Alam, *On spectral approximation of linear operators*, J. Math. Anal. Appl., 226 (1998), pp. 229–244.

[2] R. Alam and N. Behera, *Linearizations for rational matrix functions and Rosenbrock system polynomials*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 354–380.

[3] ——, *Generalized Fiedler pencils for rational matrix functions*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 587–610.

[4] A. Amparan, F. M. Dopico, S. Marcaida, and I. Zaballa, *Strong linearizations of rational matrices*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1670–1700.

[5] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, and K. Kimura, *A numerical method for nonlinear eigenvalue problems using contour integrals*, JSIAM Lett., 1 (2009), pp. 52–55.

[6] ——, *A numerical method for polynomial eigenvalue problems using contour integral*, Japan Journal of Industrial and Applied Mathematics, 27 (2010), pp. 73–90.

[7] B. Aupetit, *A primer on spectral theory*, Springer, 1991.

[8] A. P. Austin, P. Kravanja, and L. N. Trefethen, *Numerical algorithms based on analytic function values at roots of unity*, SIAM J. Numer. Anal., 52 (2014), pp. 1795–1821.

[9] J. A. Ball, I. Gohberg, and L. Rodman, *Interpolation of rational matrix functions*, vol. 45 of Operator Theory: Advances and Applications, Birkhäuser Verlag, Basel, 1990.

[10] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: a collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), pp. Art. 7, 1–28.

[11] W.-J. Beyn, *An integral method for solving nonlinear eigenvalue problems*, Linear Algebra Appl., 436 (2012), pp. 3839–3863.

[12] M. C. Brennan, M. Embree, and S. Gugercin, *Contour integral methods for nonlinear eigenvalue problems: A systems theoretic approach*, SIAM Review, 65 (2023), pp. 439–470.

[13] M. I. Bueno, F. M. Dopico, J. Pérez, R. Saavedra, and B. Zykoski, *A simplified approach to Fiedler-like pencils via block minimal bases pencils*, Linear Algebra Appl., 547 (2018), pp. 45–104.

[14] R. K. Das and R. Alam, *Affine spaces of strong linearizations for rational matrices and the recovery of eigenvectors and minimal bases*, Linear Algebra and its Applications, 569 (2019), pp. 335–368.

[15] B. Datta, *Numerical methods for linear control systems*, vol. 1, Academic Press, 2004.

[16] F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations and the recovery of minimal indices*, SIAM J. Matrix Anal. Appl., 31 (2009/10), pp. 2181–2204.

[17] L. M. Delves and J. N. Lyness, *A numerical method for locating the zeros of an analytic function*, Math. Comp., 21 (1967), pp. 543–560.

[18] M. Frost and C. Storey, *A note on the controllability of linear constant delay-differential systems*, International Journal of Control, 28 (1978), pp. 673–679.

[19] ——, *Further remarks on the controllability of linear constant delay-differential systems*, International Journal of Control, 30 (1979), pp. 863–870.

[20] T. W. Gamelin, *Complex analysis*, Springer, 2003.

[21] B. Gavin, A. Mipolhk edlar, and E. Polizzi, *FEAST eigensolver for nonlinear eigenvalue problems*, J. Comput. Sci., 27 (2018), pp. 107–117.

[22] A. Ghosh and R. Alam, *Sensitivity and backward perturbation analysis of multiparameter eigenvalue problems*, SIAM Journal on Matrix Analysis and Applications, 39 (2018), pp. 1750–1775.

[23] I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Classes of Linear Operators Vol.I*, Operator Theory: Advances and Applications, Birkhäuser Basel, 1990.

[24] I. Gohberg, M. A. Kaashoek, and F. Van Schagen, *On the local theory of regular analytic matrix functions*, Linear algebra and its applications, 182 (1993), pp. 9–25.

[25] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press Inc., New York, 1982.

[26] I. Gohberg and L. Rodman, *Analytic matrix functions with prescribed local data*, Journal D Analyse Mathematique, 40 (1981), pp. 90–128.

[27] I. C. Gohberg, M. A. Kaashoek, and D. C. Lay, *Equivalence, linearization, and decomposition of holomorphic operator functions*, J. Functional Analysis, (1978), pp. 102–144.

[28] I. C. Gohberg and E. I. Sigal, *An operator generalization of the logarithmic residue theorem and Rouché's theorem*, Mat. Sb. (N.S.), 84(126) (1971), pp. 607–629.

[29] S. Güttel and F. Tisseur, *The nonlinear eigenvalue problem*, Acta Numer., 26 (2017), pp. 1–94.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 2012.

[31] N. I. Ioakimidis and E. G. Anastasselou, *A modification of the Delves-Lyness method for locating the zeros of analytic functions*, J. Comput. Phys., 59 (1985), pp. 490–492.

[32] T. Kailath, *Linear Systems*, Prentice-Hall information and system sciences series, Prentice Hall International, 1998.

[33] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[34] M. V. Keldysh, *On the completeness of the eigenfunctions of some classes of non-selfadjoint linear operators*, Uspekhi mat. nauk, 26 (1971), pp. 15–41.

[35] V. Kozlov and V. Maz'ya, *Differential equations with operator coefficients*, Springer monographs in Mathematics, Springer, 1999.

[36] P. Kravanja and M. Van Barel, *Computing the zeros of analytic functions*, vol. 1727 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2000.

[37] P. Lancaster, *Linearization of regular matrix polynomials*, Electron. J. Linear Algebra, 17 (2008), pp. 21–27.

[38] J. Leiterer, *Local and global equivalence of meromorphic operator functions. part-ii*, Mathematische Nachrichten, 84 (1978), pp. 145–170.

[39] B. V. Limaye, *Spectral perturbation and approximation with numerical experiments*, vol. 13 of Proceedings of the Centre for Mathematical Analysis,Centre for Mathematical Analysis, Australian National University, Canberra, 1987.

[40] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Structured polynomial eigenvalue problems: good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1029–1051.

[41] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM Journal on Matrix Analysis and Applications, 28 (2006), pp. 971–1004.

[42] V. Mehrmann and C. Schröder, *Nonlinear eigenvalue and frequency response problems in industrial practice*, J. Math. Ind., 1 (2011), pp. 1–18.

[43] V. Mehrmann and H. Voss, *Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27 (2004), pp. 121–152 (2005).

[44] R. Mennicken and M. Möller, *Non-self-adjoint boundary eigenvalue problems*, vol. 192 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 2003.

[45] W. Michiels and S.-I. Niculescu, *Stability and stabilization of time-delay systems : An eigenvalue-based approach*, vol. 12 of Advances in Design and Control, SIAM, Philadelphia, PA, 2007.

[46] M. T. Nair, *On spectral properties of perturbed operators*, Proc. Amer. Math. Soc., 123 (1995), pp. 1845–1850.

[47] J. E. Osborn, *Spectral approximation for compact operators*, Math. Comp., 29 (1975), pp. 712–725.

[48] R. Remmert, *Classical topics in complex function theory*, vol. GTM.172, Springer, 1997.

[49] L. Robol, R. Vandebril, and P. Van Dooren, *A framework for structured linearizations of matrix polynomials in various bases*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 188–216.

[50] H. H. Rosenbrock, *State-Space and Multivariable Theory*, John Wiley & Sons, Inc., New York, 1970.

[51] T. Sakurai and H. Sugiura, *A projection method for generalized eigenvalue problems using numerical integration*, in Proceedings of the 6th Japan-China Joint Seminar on Numerical Mathematics (Tsukuba, 2002), vol. 159, 2003, pp. 119–128.

[52] W.-H. Steeb, *Matrix calculus and Kronecker product with applications and C++ programs*, World Scientific Publishing Co., Inc., River Edge, NJ, 1997.

[53] Y. Su and Z. Bai, *Solving rational eigenvalue problems via linearization*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 201–216.

[54] P. T. P. Tang and E. Polizzi, *Feast as a subspace iteration eigensolver accelerated by approximate spectral projection*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 354–390.

[55] M. Van Barel, *Designing rational filter functions for solving eigenvalue problems by contour integration*, Linear Algebra Appl., 502 (2016), pp. 346–365.

[56] A. I. G. Vardulakis, *Linear Multivariable Control*, John Wiley & Sons Ltd., 1991.

[57] S. Vologiannidis and E. N. Antoniou, *A permuted factors approach for the linearization of polynomial matrices*, Math. Control Signals Systems, 22 (2011), pp. 317–342.

[58] J. W.Brown and R. V. Churchill, *Complex variables and applications*, McGraw-Hill,, 2009.

[59] L. Weiss, *On the controllability of delay-differential systems*, SIAM Journal on Control, 5 (1967), pp. 575–587.

[60] S. Yokota and T. Sakurai, *A projection method for nonlinear eigenvalue problems using contour integrals*, JSIAM Lett., 5 (2013), pp. 41–44.