Diplomarbeit

Analyticity properties of the decomposition of the spectral density matrix in the context of dynamic PCA

ausgeführt am Institut für Wirtschaftsmathematik der
Technischen Universität Wien

unter der Anleitung von
O. Univ. Prof. Dipl.-Ing. Dr. tech. Manfred Deistler
durch

Bernd Funovits
Matrikelnummer 0451617
G.A. Wimmerplatz 1/2/7
7432 Oberschützen

Oberschützen am 10. August 2009 ______________________

Danksagung

"Der Realist ist insofern naiv, als er nicht zur Kenntnis nimmt, daß wir nicht in der Welt leben, sondern in dem Bild, das wir uns von der Welt machen."

Hoimar von Ditfurth

Mein Bild, das ich mir von der Welt gemacht habe, wurde durch Prof. Manfred Deistlers offene und unvoreingenommene Herangehensweise an Problemstellungen jeglicher Art wesentlich beeinflusst. Dafür möchte ich mich auf diesem Wege ganz herzlich bedanken. Seine Vorlesungen über Ökonometrie und Zeitreihen sowie das zugehörige Seminar haben für mich einen beachtlichen Mehrwert bei meinem zweijährigen Studium Aufenthalt in Frankreich geschaffen und werden mir sicherlich auch weiterhin sehr zugute kommen. Weiters schätzte ich ganz besonders, dass der Kontakt und die persönliche Betreuung auch während meiner Abwesenheit von der TU Wien niemals abriss und dass Prof. Deistler mir die Möglichkeit gab an Vortragsreihen in Brüssel und London anwesend zu sein. Sein disziplinenübergreifendes Interesse für Mathematik hat letztendlich auch zu dem Thema dieser Diplomarbeit geführt.

Ganz besonderer Dank gebührt meiner Familie die mir zu jeder Zeit Verständnis und Unterstützung für die Art und Weise wie ich zu studieren pflegte entgegenbrachte.

Contents

1 Principal component analysis in the frequency domain **4**

1.1 Introduction 4

1.1.1 Description in mathematical terms 4

1.1.2 Static and quasi-static case 5

1.1.3 Definition of principal component analysis 5

1.1.4 Example: Impulse series of earthquakes 6

1.1.5 Choice of r 6

1.2 Static case 6

1.3 Dynamic case 9

1.4 Decomposition of the spectral density 13

2 Normed commutative algebras **15**

2.1 Fundamentals 15

2.2 Maximal ideals 16

2.2.1 An example 19

2.3 Equivalence classes with respect to ideals 20

2.4 The canonical isomorphism on the maximal ideals 23

2.4.1 The relation between maximal ideals and complex multiplicative homomorphisms 24

2.5 Proof of the univariate case 27

2.6 The topology on $\mathcal{M}(R)$ 27

2.7 Generating systems and polynomially convex sets 31

2.8 Proof of the multivariate case 35

2.8.1 Case A: Algebra R is generated by a finite number of elements 35

2.8.2 General Case 39

3 Perturbation of Polynomials **42**

3.1 Polynomials with Meromorphic Coefficients 46

3.2 Behavior of the roots near critical points 52

3.3 Continuous representation of the root functions 58

4 Smoothness of Eigenprojections and Eigenvectors **61**

4.1 Some facts about resolvents and eigenprojections of matrices . . 61

4.2 Integral representation of the eigenprojections 64

4.3 Analyticity at non-critical points 66

4.3.1 Analyticity with respect to matrices 67

4.3.2 Analyticity with respect to a parameter 68

4.3.3 Analyticity and Construction of the generalized eigenbasis 72

4.4 Analytic properties of eigenprojections near critical points 74

4.5 Real Case 83

5 Bibliography **86**

1 Principal component analysis in the frequency domain

1.1 Introduction

The analysis of multivariate time series is an important issue in many research areas, such as economics, finance, signal processing and medicine. Multivariate time series are modeled jointly when the relation between the single time series or co-movements are important. In general, the number of free parameters, which is a measure of the complexity of the model class considered, increases substantially with the number of observed variables. E.g, if the cross-sectional dimension in a VAR(p) model is n, the dimension of the parameter space is proportional to n^2 (the so-called curse of dimensionality). Thus, the complexity of the model class shows quadratic dependence on n, whereas the number of data points, for fixed sample size T is linear in n. Factor models, in particular the PCA model, mitigate this problem.

1.1.1 Description in mathematical terms

The basic common equation for all different kinds of factor models considered here is of the form

$$\begin{aligned} x_t &= \Lambda(z)\xi_t + u_t \\ &= \chi_t + u_t \end{aligned} \qquad t \in \mathbb{Z},$$

where x_t is the n-dimensional vector of observations, ξ_t is the r-dimensional factor, u_t is the n-dimensional noise and $\Lambda(z) = \sum_{j=-\infty}^{\infty} \Lambda_j z^j,\ \Lambda_j \in \mathbb{R}^{n\times r}$ denotes the transfer function. Thereby, the argument z is used as complex variable as well as back-shift operator on $\mathbb{Z}$. $\chi_t = \Lambda(z)\xi_t$ is called the common component or the latent variable.

Throughout the first chapter the following is assumed:

- $\mathbb{E}(\xi_t) = \mathbb{E}(u_t) = 0, \quad \forall t \in \mathbb{Z}$
- $\mathbb{E}(\xi_t {u_s}^T) = 0 \in \mathbb{R}^{r\times n}, \quad \forall t, s \in \mathbb{Z}$
- $(\xi_t)_{t\in\mathbb{Z}}, (u_t)_{t\in\mathbb{Z}}$ are wide sense stationary and (linearly) regular

- The covariances $\gamma_\xi(s) = \mathbb{E}(\xi_t {\xi_{t+s}}^T) \in \mathbb{R}^{r\times r}$ and $\gamma_u(s) = \mathbb{E}(u_t {u_{t+s}}^T) \in \mathbb{R}^{n\times n}$ are satisfying

$$\sum_{s=-\infty}^{\infty} \|\gamma_\xi(s)\| < \infty, \quad \sum_{s=-\infty}^{\infty} \|\gamma_u(s)\| < \infty,$$

  where $\|\cdot\|$ denotes an arbitrary matrix norm.

- $\sum_{j=-\infty}^{\infty} \|\Lambda_j\| < \infty$
- The spectral density $f_\chi(\lambda)$ of $(\chi_t)_{t\in\mathbb{Z}}$ has rank r for all $\lambda \in [-\pi, \pi]$.

Due to the summability condition of $\gamma_\xi(s)$, $\gamma_u(s)$ and $\Lambda(z)$ and due to the fact that the processes ξ_t and u_t are uncorrelated, the spectral densities f_ξ of $(\xi_t)_{t\in\mathbb{Z}}$ and f_u of $(u_t)_{t\in\mathbb{Z}}$ exist as uniform limits of trigonometric polynomials and the spectral density f_x of the observation process $(x_t)_{t\in\mathbb{Z}}$ can be represented as

$$f_x(\lambda) = \Lambda(e^{-i\lambda}) f_\xi(\lambda) \overline{\Lambda(e^{-i\lambda})}^T + f_u(\lambda).$$

1.1.2 Static and quasi-static case

When $\Lambda(z) = \Lambda$ is constant and $(\xi_t)_{t\in\mathbb{Z}}$, $(u_t)_{t\in\mathbb{Z}}$ and thus $(x_t)_{t\in\mathbb{Z}}$ are white noises, the factor model is called static.

The variance matrix of $(x_t)_{t\in\mathbb{Z}}$, $\Sigma_x \in \mathbb{R}^{n\times n}$, is decomposed as

$$\Sigma_x = \Lambda \Sigma_\xi \Lambda^T + \Sigma_u.$$

When $(\xi_t)_{t\in\mathbb{Z}}$, $(u_t)_{t\in\mathbb{Z}}$ are not necessarily white noises, the model is called quasi-static.

For given f_x, or Σ_x respectively, too many models would be possible, see [14]. Thus, in order to obtain reasonable model classes, further assumptions have to be imposed. This leads to principal component models, linear factor models with idiosyncratic noise and generalized linear factor model. In this work, only the first-mentioned will be considered.

1.1.3 Definition of principal component analysis

The aim of principal component analysis (PCA) is to approximate the n-dimensional observed process $(x_t)_{t\in\mathbb{Z}}$ by a filtered version of itself whose spectral density is of reduced rank r such that the variance of the residuals is minimized. Hence, the additional assumption in the PCA-model is, that $\xi_t = C(z)x_t$ holds. The $(r \times n)$-dimensional filter $C(z)$ and the $(n \times r)$-dimensional filter $\Lambda(z)$ are obtained by minimizing for fixed reduced rank r

$$\operatorname{tr}\left\{\mathbb{E}(u_t {u_t}^T)\right\} = \operatorname{tr}\left\{\mathbb{E}[(x_t - \Lambda(z)C(z)x_t)(x_t - \Lambda(z)C(z)x_t)^T]\right\},$$

where tr denotes the trace.

1.1.4 Example: Impulse series of earthquakes

A situation in which such a model of reduced rank might be employed is the following: Let ξ_t represent the impulse series of r earthquakes occurring simultaneously at various locations; let x_t represent the signals received by n seismometers; let $\Lambda(z)$ represent the transmission effects of the earth on the earthquakes. Seismologists are interested in investigating the series $(\xi_t)_{t \in \mathbb{Z}}$.

1.1.5 Choice of r

For the PCA - model, the number of factors r is not intrinsic in the sense, that it is not a property of f_x or Σ_x. By the choice of r, the degree of dimension reduction and, as a trade-off, the quality of approximation are determined.

1.2 Static case

As already mentioned above, in the static case of the principal component analysis $\Lambda(z) = \Lambda \in \mathbb{R}^{n \times r}$ and $C(z) = C \in \mathbb{R}^{r \times n}$ are constant.

The solution of the minimization problem

$$\min_{\substack{\Lambda \in \mathbb{R}^{n \times r} \\ C \in \mathbb{R}^{r \times n}}} \operatorname{tr}\left\{\mathbb{E}[(x_t - \Lambda C x_t)(x_t - \Lambda C x_t)^T]\right\},$$

is obtained by the eigenvalue decomposition of the variance-covariance matrix of the random variable x_t.

The following theorem treats the case where $\mathbb{E}(x) = c \neq 0 \in \mathbb{R}^n$.

Theorem 1.1 (Solution of the static minimization problem)**.**
Let x_t be a n-dimensional vector-valued random variable with the properties

- $\mathbb{E}(x_t) = c_x \in \mathbb{R}^n, \quad \forall t \in \mathbb{Z}$
- $\mathbb{E}[(x_t - c_x)(x_t - c_x)^T] = \Sigma_x$.

Then, the solution of the minimization problem

$$\min_{\substack{\Lambda \in \mathbb{R}^{n\times r}\\ C \in \mathbb{R}^{r\times n}\\ d \in \mathbb{R}^n}} \operatorname{tr}\left\{\mathbb{E}[(x_t - (d + \Lambda C x_t))(x_t - (d + \Lambda C x_t))^T]\right\}$$

is given by

- $d = c_x - \Lambda C c_x$
- $C = \begin{pmatrix} \dots & {o_1}^T & \dots \\ \dots & {o_2}^T & \dots \\ \dots & \vdots & \dots \\ \dots & {o_r}^T & \dots \end{pmatrix} =: {O_1}^T, \qquad \Lambda = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ o_1 & o_2 & \dots & o_r \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} = O_1,$

where the o_j are the eigenvectors of Σ_x.

The obtained minimal value is

$$\min_{\substack{\Lambda \in \mathbb{R}^{n\times r}\\ C \in \mathbb{R}^{r\times n}\\ d \in \mathbb{R}^n}} \operatorname{tr}\left\{\mathbb{E}((x_t - (d + \Lambda C x_t))\,(x_t - (d + \Lambda C x_t))^T)\right\} = \sum_{j>r} \mu_j,$$

where the μ_j are the eigenvalues of Σ_x and $\mu_{j+1} \le \mu_j$ for all $j \in \{1, \dots, n-1\}$

.

Proof. See [2]. □

Post-multiplying Λ with a non-singular matrix $P \in \mathbb{R}^{n\times n}$ and pre-multiplying ξ_t with its inverse P^{-1} yields the same χ_t. However, by making the special choice

$$\xi_t = {O_1}^T x_t, \quad \Lambda = O_1, \quad u_t = O_2 {O_2}^T x_t$$

the matrices in the decomposition of Σ_x

$$\begin{aligned} \Sigma_x &= O_1 \Omega_1 {O_1}^T + O_2 \Omega_2 {O_2}^T = \\ &= \Sigma_\chi + \Sigma_u, \end{aligned}$$

can be uniquely identified. In this representation, $\Omega_1 \in \mathbb{R}^{r\times r}$ and $\Omega_2 \in \mathbb{R}^{(n-r)\times(n-r)}$ denote the diagonal matrices consisting of the r largest and $(n-r)$

smallest eigenvalues of Σ_x arranged in decreasing order and $O_1 = (o_1 \cdots o_r)$ and $O_2 = (o_{r+1} \cdots o_n)$ the $(n \times r)$- and $[n \times (n-r)]$-dimensional matrices of the corresponding eigenvectors.

Definition 1.2 (Principal component ξ_t of x_t). *The variate* $\xi_t^j = {o_j}^T x_t, j \in \{1, \ldots, n\}, t \in \mathbb{Z}$ *is called the j-th principal component of* $(x_t)_{t \in \mathbb{Z}}$.

- The principal components can be interpreted as the coordinates of x_t in the eigenbasis of Σ_x.

If x_t follows a normal distribution with n dimensions, i.e. $N_n(0, \Sigma_x)$, the components ξ_t^j of the vector of principal components are independent and $N_1(0, \mu_j)$-distributed, where μ_j denotes the j-th largest eigenvalue of Σ_x.

Let $X_m, m \in \{1, \ldots, T\}$ be a sample from $N_n(0, \Sigma_x)$ and define $\mathbf{X} = [X_1 \cdots X_T] \in \mathbb{R}^{n \times T}$. An estimator for Σ_x is

$$\hat{\Sigma}_x = \frac{\mathbf{X}\mathbf{X}^T}{T}.$$

The j-th largest eigenvalue μ_j of Σ_x is estimated by the j-th largest eigenvalue $\hat{\mu}_j$ of $\hat{\Sigma}_x$ and the corresponding eigenvector o_j by $\hat{o}_j$.

Theorem 1.3 (Results for estimators). *Under the assumption that*

- $X_1, \ldots, X_T$ *is a sample from* $N_n(0, \Sigma_x)$ *and that*
- *the eigenvalues of* Σ_x *are distinct,*

it follows, that

- *the members of the set* $\{\hat{\mu}_j, \hat{o}_j : j \in \{1, \ldots, n\}\}$ *are asymptotically normal and*
- $\{\hat{\mu}_j, : j \in \{1, \ldots, n\}\}$ *and* $\{\hat{o}_j : j \in \{1, \ldots, n\}\}$ *are asymptotically independent.*

Furthermore, the asymptotic moments are given by

- $\mathbb{E}(\hat{\mu}_j) = \mu_j + \mathcal{O}(T^{-1})$

- $\mathbb{E}(\hat{o}_j) = o_j + \mathcal{O}(T^{-1})$
- $Cov(\hat{\mu}_j, \hat{\mu}_k) = \begin{cases} \delta_{jk}\frac{\mu_j^2}{T} + \mathcal{O}(T^{-2}) & j = k \\ \mathcal{O}(T^{-2}) & j \neq k \end{cases}$
- $Cov(\hat{o}_j, \hat{o}_k) = \begin{cases} \mu_j \sum\limits_{l \neq j} \frac{\mu_l}{(\mu_j - \mu_l)^2} \frac{o_l {o_l}^T}{T} + \mathcal{O}(T^{-2}) & j = k \\ \mathcal{O}(T^{-2}) & j \neq k \end{cases}$
- $Cov(\hat{o}_j, \hat{o}_k) = \begin{cases} -\frac{\mu_k \mu_j}{(\mu_j - \mu_l)^2} \frac{o_k o_j^T}{T} + \mathcal{O}(T^{-2}) & j = k \\ \mathcal{O}(T^{-2}) & j \neq k \end{cases}$

where $j, k \in \{1, \ldots, n\}$.

Proof. See [2]. □

The theorem above results mainly from two facts:

- The indicated eigenvalues and vectors are differentiable functions of the entries of $\hat{\Sigma}_x$.
- $\hat{\Sigma}_x$ is asymptotically normal as $T \to \infty$.

The matrix $\hat{\Sigma}_x$ has a Wishart distribution, which is a multidimensional generalization of the χ^2-distribution, and the distribution of $\hat{\mu}_j$ can be approximated by $\frac{\mu_j \chi^2_{2T}}{2T}$.

1.3 Dynamic case

For the dynamic generalization of the PCA, the spectral density f_x and its eigenvalue decomposition are considered in a similar way. The filter $\widetilde{O_1(z)}$ of reduced rank which provides the best approximation of the process by itself will be deduced.

The basic equation is again

$$x_t = \widetilde{O_1(z)}\xi_t + u_t = \chi_t + u_t, \quad t \in \mathbb{Z}.$$

This matrix function can, due to the fact that $(u_t)_{t\in\mathbb{Z}}$ and $(\xi_t)_{t\in\mathbb{Z}}$ are uncorrelated, be decomposed as

$$f_x(\lambda) = \widetilde{O_1}(e^{-i\lambda})\Lambda_1(\lambda)\overline{\widetilde{O_1}(e^{-i\lambda})}^T + \widetilde{O_2}(e^{-i\lambda})\Lambda_2(\lambda)\overline{\widetilde{O_2}(e^{-i\lambda})}^T, \tag{1}$$

where $\widetilde{O_1}(e^{i\lambda})$ and $\widetilde{O_2}(e^{i\lambda})$ are the corresponding transfer function to the $(n\times r)$- and $[n\times(n-r)]$-dimensional filters $\widetilde{O_1}(z)$ and $\widetilde{O_2}(z)$. For the sake of simplicity, the transfer functions are, from now on, denoted as $O_1(\lambda)$ and $O_2(\lambda)$.

Theorem 1.4 (Solution of the minimization problem)**.** *Let $(x_t)_{t\in\mathbb{Z}}$ be a n-dimensional wide sense stationary stochastic process with absolutely summable covariance function satisfying the conditions*

- $\mathbb{E}(x_t) = c_x \quad \forall t \in \mathbb{Z}$,
- $\operatorname{Cov}(x_{t+u}, x_t) = \gamma_x(u) \in \mathbb{R}^{n\times n}, u \in \mathbb{Z}$ *is absolutely summable and*
- $f_x(\lambda) = \frac{1}{2\pi}\sum\limits_{u=-\infty}^{\infty}\gamma_x(u)e^{-iu\lambda} \in \mathbb{C}^{r\times r}$ *is positive-definite for every* $\lambda \in \mathbb{R}$.

The solution of the minimization problem

$$\min_{\substack{\Lambda(z)\\ C(z)\\ d\in\mathbb{R}^n}} \operatorname{tr}\left\{\mathbb{E}((x_t - (d+\Lambda(z)C(z)x_t))\,(x_t-(d+\Lambda(z)C(z)x_t))^T)\right\}, \tag{2}$$

where

- $\Lambda(z) = \sum\limits_{j=-\infty}^{\infty}\Lambda_j z^j, \Lambda_j \in \mathbb{R}^{n\times r}$ *is a $(n\times r)$-dimensional filter and*
- $C(z) = \sum\limits_{j=-\infty}^{\infty} C_j z^j, C_j \in \mathbb{R}^{r\times n}$ *is a $(r\times n)$-dimensional filter,*

is given by

- $d = c_x - \Lambda(z)C(z)c_x$
- $C(\lambda) = \begin{pmatrix} \dots & \overline{o_1(\lambda)}^T & \dots \\ \dots & \overline{o_2(\lambda)}^T & \dots \\ \dots & \vdots & \dots \\ \dots & \overline{o_r(\lambda)}^T & \dots \end{pmatrix}, \Lambda(\lambda) = \begin{pmatrix} \vdots & \vdots & \vdots \\ o_1(\lambda) & \dots & o_r(\lambda) \\ \vdots & \vdots & \vdots \end{pmatrix} = \overline{C}^T(\lambda)$
- $\Lambda_j = \int\limits_0^{2\pi}\Lambda(\lambda)e^{ij\alpha}d\alpha, \quad C_j = \int\limits_0^{2\pi} C(\lambda)e^{ij\alpha}d\alpha,$

where the $o_j(\lambda)$ are the eigenvectors of $f_x(\lambda)$.

The obtained minimal value is

$$\int_0^{2\pi}\sum_{j>r}\mu_j(\alpha)d\alpha,$$

where $\mu_j(\lambda)$ is the j-th eigenvalues of $f_x(\lambda)$.

Proof. Step 1: Reformulation of the minimization problem.
Since

- $\mathbb{E}(Y^T \cdot Y) = \mathbb{E}(Y - \mathbb{E}(Y))^T \cdot (Y - \mathbb{E}(Y)) + \mathbb{E}(Y)^T \mathbb{E}(Y)$ and
- $\operatorname{tr}\left\{\mathbb{E}(YY^T)\right\} = \mathbb{E}(\operatorname{tr}\left\{YY^T\right\}) = \mathbb{E}(\operatorname{tr}\left\{Y^TY\right\}) = \operatorname{tr}\left\{\mathbb{E}(Y^TY)\right\} = \mathbb{E}(Y^TY)$,

where Y is a n-dimensional random variable and by defining

$$\begin{aligned}\varepsilon_t &:= x_t - (d + \Lambda(z)C(z)x_t) - \mathbb{E}(x_t - (d + \Lambda(z)C(z)x_t)) \\ &= x_t - (d + \Lambda(z)C(z)x_t) - c_x + d + \Lambda(z)C(z)c_x,\end{aligned}$$

the expression

$$\operatorname{tr}\left\{\mathbb{E}((x_t - (d + \Lambda(z)C(z)x_t))\,(x_t - (d + \Lambda(z)C(z)x_t))^T)\right\}$$

is minimal if and only if

$$\mathbb{E}[(x_t - d - \Lambda(z)C(z)x_t)^T]\mathbb{E}[(x_t - d - \Lambda(z)C(z)x_t)] + \mathbb{E}({\varepsilon_t}^T\varepsilon_t) \qquad (3)$$

is minimal.

Step 2a: Minimization of the first term.
The first term in (3) is annihilated by setting

$$d := c_x - \Lambda(z)C(z)c_x.$$

Step 2b: Minimization of the second term.
The second term in (3) can be written as

$$\begin{aligned}\mathbb{E}({\varepsilon_t}^T\varepsilon_t) &= \mathbb{E}\left(\operatorname{tr}\left\{\varepsilon_t{\varepsilon_t}^T\right\}\right) = \\ &= \operatorname{tr}\left\{\mathbb{E}((\varepsilon_t{\varepsilon_t}^T))\right\} = \\ &= \operatorname{tr}\left\{\gamma_\varepsilon(0)\right\} \\ &= \operatorname{tr}\left\{\int_0^{2\pi} e^{i0\lambda} f_\varepsilon(\lambda)d\lambda\right\}\end{aligned}$$

By defining $A(\lambda) := \Lambda(e^{i\lambda})C(e^{i\lambda})$

$$f_\varepsilon(\lambda) = (I - A(\lambda))f_x(\lambda)\overline{(I - A(\lambda))}^T.$$

Since $f_x(\lambda)$ is positive-definite $f_x(\lambda)^{\frac{1}{2}} = O(\lambda) \cdot \sqrt{\Omega(\lambda)} \cdot \overline{O(\lambda)}^T$ exists, $f_x(\lambda) = \overline{f_x(\lambda)}^T$ and therefore

$$\begin{aligned} &\operatorname{tr}\left[(I - A(\lambda)) f_x(\lambda) \overline{(I - A(\lambda))}^T\right] = \\ &= \operatorname{tr}\left[(f_x(\lambda)^{\frac{1}{2}} - A(\lambda) f_x(\lambda)^{\frac{1}{2}}) \overline{(f_x(\lambda)^{\frac{1}{2}} - A(\lambda) f_x(\lambda)^{\frac{1}{2}})}^T\right] \end{aligned}$$

is minimized over all transfer function $A(\lambda)$ with rank smaller or equal to r by

$$A(\lambda) = \sum_{j=1}^{r} o_j(\lambda) \overline{o_j(\lambda)}^T.$$

The minimal value of (2) is

$$\begin{aligned} &\operatorname{tr}\left[\underbrace{(O(\lambda)\sqrt{\Omega(\lambda)}\overline{O(\lambda)}^T - A(\lambda)O(\lambda)\sqrt{\Omega(\lambda)}\overline{O(\lambda)}^T)}_{=*} \overline{(*)}^T\right] = \\ &= \operatorname{tr}\left[\underbrace{(O(\lambda)\sqrt{\Omega(\lambda)}\overline{O(\lambda)}^T - \sum_{j=1}^{r} o_j(\lambda)\overline{o_j(\lambda)}^T O(\lambda)\sqrt{\Omega(\lambda)}\overline{O(\lambda)}^T)}_{=*} \overline{(*)}^T\right] = \\ &tr(\Omega_{n-r}(\lambda)), \end{aligned}$$

where $\Omega_{n-r}(\lambda)$ denotes the $(n-r) \times (n-r)$ diagonal matrix containing those $(n-r)$ smallest eigenvalues $\square$

1.4 Decomposition of the spectral density

By considering the expression (1) on page 9, many interesting questions arise.

- Do all the eigenvalues of $f_x(\lambda)$ have to be distinct that they can represented as analytic or continuous functions?
    - If they are distinct, it follows from the implicit functions theorem that they can be represented, under certain assumptions, as differentiable functions. However, the question will be answered positively by a theorem of Kato [11].
- If the coefficients of $f_x(\lambda)$ satisfy a summability condition which permit to state that stationary input creates stationary output, is this summability condition also fulfilled by the transfer function $\overline{O_1(\lambda)}^T$?

Every element in the transfer function $\overline{O_1(\lambda)}^T$ is an analytical transformation of the coefficients of the spectral density. Given that the coefficients of the spectral density satisfy a summability condition (which will be stated in the following theorem), it has to be shown that the analytical transformation of the coefficients of $f_x(\lambda)$ satisfy the same summability condition in order to prove the aforementioned question positively.

The miscellaneous summability conditions on the coefficients of the spectral density are stated by the means of the normed commutative Banach algebra

$$V(l) = \left\{ z(\lambda) = \sum_{u=-\infty}^{\infty} a(u) e^{-iu\lambda}, \lambda \in \mathbb{R} \mid a(u) \in \mathbb{C}\ \forall u \in \mathbb{Z} \wedge \sum_{u=-\infty}^{\infty} (1 + |u|^l)|a(u)| < \infty \right\}.$$

Theorem 1.5 (Transformation of multivariate time series)**.** *Under the assumptions*

- *that the elements* $f_x^{(i,j)}(\lambda)$ *of* $f_x(\lambda)$ *are elements of* $V(p)$*, i.e.* $\sum_{u=-\infty}^{\infty}(1 + |u|^p)|\gamma_x^{(i,j)}(u)| < \infty \quad \forall i, j \in \{1, \dots, n\}$*, where* $\gamma_x^{(i,j)}(u)$ *denotes the* (i, j)*-element of the covariance function of* $(x_t)_{t \in \mathbb{Z}}$*,*
- *that* $f_x(\lambda)$ *is self-adjoint, i.e.* $f_x(\lambda) = \overline{f_x(\lambda)}^T \quad \forall \lambda \in (-\pi, \pi]$*, and* l *is its maximum number of distinct eigenvalues and*

- *that* $f_x(\lambda)$ *can be continued analytically in* $\mathbb{C}$ *across a neighborhood which contains the set* $\left\{ \left(f^{(i,j)}(\lambda) \right)_{i,j=1}^{n} \mid \lambda \in (-\pi, \pi] \right\}$ *and that it is there self-adjoint as well.*

it follows that

- *the transfer functions* $O_1(\lambda)$ *and* $\overline{O_1(\lambda)}^T$ *are in* $V(p)$ *as well, i.e.*
  - *for every coefficient* $O_1^{(i,j)}(\lambda)$ *of* $O_1(\lambda)$ *exists a sequence* $(b^{(i,j)}(u))_{u\in\mathbb{Z}}$*, such that*

    $$O_1^{(i,j)}(\lambda) = \textstyle\sum_{u=-\infty}^{\infty} b^{(i,j)}(u) e^{-iu\lambda}$$

    and

    $$\textstyle\sum_{u=-\infty}^{\infty} (1 + |u|^p) |b^{(i,j)}(u)| < \infty$$

    and
  - *for every coefficient* $\overline{O_1^{(i,j)}(\lambda)}^T$ *of* $\overline{O_1(\lambda)}^T$ *exists a sequence* $(c^{(i,j)}(u))_{u\in\mathbb{Z}}$*, such that*

    $$\overline{O_1^{(i,j)}(\lambda)}^T = \sum_{u=-\infty}^{\infty} c^{(i,j)}(u) e^{-iu\lambda}$$

    and

    $$\sum_{u=-\infty}^{\infty} (1 + |u|^p) |c^{(i,j)}(u)| < \infty.$$

Furthermore, the spectral density matrices $f_\xi(\lambda)$ *and* f_ε *of* $(\xi)_{t\in\mathbb{Z}}$ *and* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *respectively are in* $V(p)$ *too.*

Proof. By corollary 4.21, the coefficients of the transfer functions $O_1(\lambda)$ and $\overline{O_1(\lambda)}^T$ are analytic functions of the entries of $f_x(\lambda)$.

Therefore, by lemma 2.37, the summability condition is satisfied by the entries of $O_1(\lambda)$ and $\overline{O_1(\lambda)}^T$ as well.

Furthermore, the spectral density matrices $f_\xi(\lambda)$ and f_ε of $(\xi)_{t\in\mathbb{Z}}$ and $(\varepsilon_t)_{t\in\mathbb{Z}}$ respectively are in $V(p)$ too, because $\xi_t = \widetilde{O_1(z)} x_t$ and $\varepsilon_t = \widetilde{O_2(z)} x_t$.

□

2 Normed commutative algebras

In this chapter, it will be proved that the analytic transformation of several elements of an commutative normed algebra is an element of the algebra as well. This result, corollary 2.1, is then applied to the elements of the spectral density matrix $f_x(\lambda)$ which are contained within $V(l)$ of a stochastic process $(x_t)_{t\in\mathbb{Z}}$ with the properties stated in the preceding chapter. The elements of $f_x(\lambda)$'s eigenbasis are as analytic transformation of elements of $V(l)$ again contained within $V(l)$. The fact that the eigenbasis depends analytically on the entries of $f_x(\lambda)$ and therefore on $\lambda \in (-\pi, \pi]$ will be proved in the last chapter.

Firstly, the result will be proved for the case, where only one member of the algebra is transformed analytically. Secondly, the theorem will be proved for the multidimensional case. Except for some explicitly emphasized facts, the whole section is based on [6].

As already mentioned above, the following corollary is the main result of this chapter. It follows directly from theorem 2.37.

Corollary 2.1. *If $z_j(\lambda)$ belongs to $V(l)$, $j \in \{1, \dots, n\}$, and $f(\zeta_1, \dots, \zeta_n)$ is analytic on a neighborhood of the range of values $\{(z_1(\lambda), \dots, z_n(\lambda)) \mid \lambda \in (-\pi, \pi]\}$, then $f(z_1(\lambda), \dots, z_n(\lambda))$ also belongs to $V(l)$.*

In order to prove the univariate case, some theory about maximal ideals, their relation to complex multiplicative homomorphisms on the algebra and their quotient algebras must be developed.

2.1 Fundamentals

Definition 2.2 (Normed commutative algebra R)**.** *A normed commutative algebra R is a complex Banach space with an associative, commutative and left- and right-continuous multiplication*

$$(\cdot, \cdot) : \begin{cases} R \times R & \to \quad R \\ (x, y) & \mapsto \quad x \cdot y \end{cases} .$$

For every normed algebra without neutral element e for multiplication, a formal neutral element can be adjoined. One defines the algebra of the formal sums

$\lambda e + x$, $\lambda \in \mathbb{C}$ and $x \in R$. Therefore, it is assumed without loss of generality that the normed commutative algebra have a neutral element for multiplication.

Theorem 2.3 (Norm-inequality)**.** *For every normed algebra R with neutral element $e \neq 0$, there exists a norm on R which induces the same topology as the initial one and which satisfies*

$$\|x \cdot y\| \leq \|x\|\|y\| \quad \forall\, x, y \in R \text{ and} \tag{4}$$

$$\|e\| = 1. \tag{5}$$

Proof. See [13], theorem 10.2. □

Because of the preceding theorem, it is assumed that the norm $\|\cdot\|$ of the algebra R satisfies the condition

$$\|x \cdot y\| \leq \|x\|\|y\|, \quad \forall x, y \in R.$$

Definition 2.4 (Absolutely convergent series)**.** *All the x_i are contained within the algebra R. The series $\lim_{N\to\infty} \sum_{i=1}^{N} x_i$ is called absolutely convergent if and only if the series*

$$\lim_{N\to\infty} \sum_{i=1}^{N} \|x_i\|$$

converges.

A neighborhood of an element $x \in R$ with radius ε is denoted by

$$U_\varepsilon(x) := \{y \in R \mid \|x - y\| < \varepsilon\}$$

The collection of all the neighborhoods of $x \in R$ is called the neighborhood system. The neighborhood system is denoted by $\mathcal{U}(x)$.

Furthermore, a superscript indicates the underlying space, if it is not evident.

2.2 Maximal ideals

Theorem 2.5 (The set of inverse elements is open)**.**

- $O = \left\{x \in R \mid \exists\ x^{-1}\right\}$ *is an open set in the norm-topology.*

- $(\cdot)^{-1} : \begin{cases} O & \rightarrow & R \\ x & \mapsto & x^{-1} \end{cases}$ *is a continous map from O on to R.*

Proof. STEP 1A: IF THE DISTANCE BETWEEN $x \in R$ AND THE NEUTRAL ELEMENT e IS SMALLER THAN 1, THE INVERSE ELEMENT EXISTS AND CAN BE REPRESENTED AS SERIES, I.E. $\|\mathbf{e}-\mathbf{x}\| < \mathbf{1} \Longrightarrow \exists \mathbf{x}^{-\mathbf{1}}, \sum_{n=0}^{\infty}(e-x)^n = x^{-1}$
The series $e+(e-x)+(e-x)^2+\cdots$ converges absolutely because $\|(e-x)^n\| \leq \underbrace{\|e-x\|^n}_{<1}$. Since R is complete, this sum is also an element of R.

Multiplying this series with $x = e-(e-x)$ leads to the telescoping sum

$$\sum_{n=0}^{\infty}(e-x)^n \cdot \underbrace{(e-(e-x))}_{=x} = e+(e-x)+(e-x)^2+\cdots-(e-x)-(e-x)^2-\cdots = e.$$

Thus, the series is the inverse element of x.

STEP 1B: O IS OPEN, I.E. $\forall \mathbf{x} \in \mathbf{O} \Longrightarrow \exists \mathbf{U}(\mathbf{x}) \in \mathcal{U}(\mathbf{x}) : \mathbf{U}(\mathbf{x}) \subseteq \mathbf{O}$.
Due to the paragraph above, $U_1(e) = \{y \in R : \|e-y\| < 1\}$ is a subset of O. It follows from the continuity of the multiplication and the equation $x \cdot x^{-1} = e$ that there exists a neighborhood $U_{\delta(1)}(x)$ with $U_{\delta(1)}(x)x^{-1} \subseteq U_1(e)$. This fact can be expressed as

$$\forall z \in U_{\delta(1)}(x) : \exists (zx^{-1})^{-1}, \text{i.e. } zx^{-1}(zx^{-1})^{-1} = e.$$

It is easy to see that $x^{-1}(zx^{-1})^{-1}$ is the inverse element of z. Therefore, the statement above is proved.

STEP 2: CONTINUITY OF THE INVERSE FUNCTION, I.E. $\mathbf{x}_\mathbf{n}^{-\mathbf{1}} \rightarrow \mathbf{x}^{-\mathbf{1}}$
The series $(z_n)_{n\in\mathbb{N}}$, defined by $z_n = x_n x^{-1}$ converges to e, whenever x_n converges to x. Therefore, z_n^{-1} exists for all n with $\|z_n - e\| < 1$. Furthermore,

$$z_n^{-1} = e + (e-z_n) + (e-z_n)^2 + \cdots$$

and converges to e, whenever x_n converges to x.

It is straightforward that the application $(\cdot)^{-1}$ is continuous on O.

$$\mathbf{x}_\mathbf{n}^{-\mathbf{1}} = x^{-1}xx_n^{-1} = x^{-1}z_n^{-1} \xrightarrow{n\to\infty} x^{-1}e = \mathbf{x}^{-\mathbf{1}}$$

□

Definition 2.6 ((Proper) ideal). *A subset I of the algebra R is called ideal if the following conditions are satisfied.*

- *If x and y are contained within I, then their sum $x + y$ is an element of I as well, i.e. $x, y \in I \Rightarrow x + y \in I$.*
- *If x is an element of I and z an element of the algebra R, then $x \cdot z$ is an element of I as well, i.e. $x \in I \Rightarrow zx \in I \quad \forall z \in R$.*

If in addition $I \neq R$, the ideal I is called a proper ideal of R.

Theorem 2.7 (Relation between inverse elements and proper ideals). *An element x of R has an inverse element $x^{-1} \in R$ if and only if there does not exist a proper ideal of R which contains x, i.e.*

$$\exists\, x^{-1} \iff x \notin I_p, \quad \forall \text{ proper ideals } I_p \text{ of } R.$$

Proof. **"⇒":** If the inverse x^{-1} exists, x can not be contained within a *proper* ideal because $\underbrace{(z \cdot x^{-1})}_{\in R} \cdot x \in I$. Therefore, I coincides with R.

"⇐": If x^{-1} does not exist, $I = \{z \cdot x | z \in R\}$ is a proper ideal because the neutral element e is not contained within I.

□

Corollary 2.8 (Closure of a proper ideal). *The closure $\overline{I}$ of a proper ideal I is a proper ideal as well.*

Proof. Due to theorem 2.5 and theorem 2.7, every proper ideal is contained within $R \backslash O$, which is as complement of on open set closed. Therefore, $\overline{I}$ belongs to $R \backslash O$ as well. □

Definition 2.9 (Maximal ideal). *A maximal ideal M is a proper ideal which is not contained within any other proper ideal of R.*

Theorem 2.10 (Properties of maximal ideals).

- *Every proper ideal I of R is contained within a maximal ideal of R.*

- *Every maximal ideal M of R is closed.*

Proof. Step 1: The first part is proved by transfinite induction and is omitted, see [13] theorem 11.3.
Step 2: As proper ideal the closure $\overline{M}$ of the maximal ideal M is also a proper ideal. Since M does not belong to any larger proper ideal it must coincide with $\overline{M}$. ☐

The following theorem concludes about the structure of a normed commutative ring in terms of its maximal ideals.

Theorem 2.11 (Relation between inverse elements and maximal ideals). *An element x of R has an inverse element $x^{-1} \in R$ if and only if there does not exist a maximal ideal of R which contains x, i.e.*

$$\exists\, x^{-1} \iff x \notin M, \quad \forall \text{ maximal ideals } M \text{ of } R$$

If the only maximal ideal of R is the trivial one, it follows that R is a field.

Proof. The statement follows directly from theorem 2.7 and theorem 2.10. ☐

2.2.1 An example

Let M_τ be the set of all functions of $V(l)$ which vanish in the point τ. Then, M_τ is a maximal ideal in $V(l)$. The same statement holds for any algebra of functions.

The domain S of the functions in $V(l)$ and its maximal ideals M_τ are respectively defined by

$$\begin{aligned} &S := \{\lambda \in \mathbb{R} \mid -\pi < \lambda \leq \pi\} \text{ and} \\ &M_\tau = \{z(\lambda) \in V(l) \mid \exists \tau \in S : z(\tau) = 0\}\,. \end{aligned}$$

M_τ is a maximal ideal:

It is obvious that M_τ is a proper ideal in R. Furthermore, every function $z(\lambda)$ in $V(l)$ can be represented as

$$z(\lambda) = \frac{z(\tau)}{y(\tau)}y(\lambda) + \left(z(\lambda) - \frac{y(\lambda)}{y(\tau)}z(\tau)\right).$$

The second summand is a member of M_τ. The first summand is a multiple of $y(\lambda) \notin M_\tau$. Hence, a proper ideal which contains M_τ and $y(\lambda) \notin M_\tau$ does not exist. □

2.3 Equivalence classes with respect to ideals

Definition 2.12 (Congruence with respect to an ideal I)**.** *Two elements x and y of a commutative normed algebra R are congruent with respect to an ideal I if and only if their difference is in I, i.e.*

$$x, y \in R : x \sim y \iff x - y \in I.$$

Since this relation is reflexive, symmetric and transitive, R is partitioned in classes of congruent elements. The quotient algebra of R with respect to I is denoted by R/I. Consider for example the quotient algebra of continuous functions on a compact set with respect to a maximal ideal, e.g. the maximal ideal containing all the functions vanishing in one point. The equivalence classes consist of those function which are equal in τ.

In the following, the equivalence class pertaining to x is denoted by $[x]_\sim$. By contrast, small letters are used for the elements of the algebra.

Theorem 2.13 (R/I is a normed commutative algebra)**.** *Let R be a commutative normed algebra. The norm on R/I is defined by*

$$\|[x]_\sim\|_{R/I} = \inf_{x \in [x]_\sim} \|x\|_R.$$

If I is a closed proper ideal, then R/I is a normed commutative algebra as well.

Proof. It is known from the Banach space theory that $\inf\limits_{x \in [x]_\sim} \|x\|$ is a norm on the quotient space of the Banach space R, see [13]. The multiplicativity of the norm is evident as well.

Only the additional condition about the norm of the neutral element $[e]_\sim$ in Banach algebras will be proved.

$\|\mathbf{E}\|_{\mathbf{R/I}} \leq \mathbf{1}$: Since e is in $[e]_\sim$, the norm of $[e]_\sim$ has to be at most 1.

$\|\mathbf{E}\|_{\mathbf{R/I}} \geq \mathbf{1}$: If y is contained within $[e]_\sim$, an element $x \in I$ exists which satisfies $y = e + x$. This is because $x = y - e$ and because y and e are in the same equivalence class $[e]_\sim$. If $\|y\|_R$ were strictly smaller than 1, the inverse element $(y-e)^{-1} = x^{-1}$ of $(y-e)$ would exist. This is a contradiction to the fact that x is a member of the proper ideal I.

□

Theorem 2.14. *The homomorphism H_I*

$$H_I : \begin{cases} R & \rightarrow & R/I \\ x & \mapsto & [x]_\sim \end{cases}$$

assigns to every element $x \in R$ the equivalence class $[x]_\sim$ with respect to I. For every closed ideal I, H_I is an open and continuous application.

Proof. Step 1: H_I is open.
Let $U_\delta(0) = \{x \in R \mid \|x\|_R < \delta\} \subseteq R$ be an open ball cent red at zero. Due to the definition of the norm on R/I, the image $U' = \{H_I(x) \in R/I \ : \ x \in U_\delta(0)\}$ of $U_\delta(0)$ in R/I consists of all equivalence classes with $\|[x]_\sim\| < \delta$. Thus, the image U' is an open subset in R/I.

This means in mathematical terms that

$$\|H_I(x)\|_{R/I} = \|[x]_\sim\|_{R/I} = \inf_{x \in [x]_\sim} \|x\|_R \leq \|x\|_R < \delta.$$

Since the open balls constitute a basis of the topology of R, the image of every open subset in R is an open subset in R/I.

Step 2: H_I is continuous, i.e. if F' is closed in R/I, then the preimage $F = H_I^{-1}(F')$ is closed in R as well.
Let $(x_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in F with limit x in R. Furthermore, x_i and the limit point x are respectively contained within the corresponding equivalences $[x_i]_\sim$ and $[x]_\sim$. It must be shown that $\lim_{n\to\infty} x_n = x$ is contained within F.

From the inequality $\|[x]_\sim - [x_n]_\sim\| = \inf_{(x-x_n) \in [x-x_n]_\sim} \|x - x_n\| \leq \|x - x_n\|$, it follows that the corresponding equivalence classes $[x_n]_\sim$ satisfy

$$\lim_{n\to\infty} [x_n]_\sim = [x]_\sim.$$

Furthermore, $[x]_\sim$ is contained within $F' \subseteq R/I$, because F' is closed. Now, the arbitrarily chosen Cauchy sequence $(x_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$ converges towards an element $x \in F$. Therefore, F is a closed subset in R. □

Theorem 2.15. *For $J \supseteq I$, there exists a bijective relation between the closed ideals J of R and J' of R/I.*

Proof. Step 1: If J' closed, then J closed and H_I^{-1} is injective.
Since H_I is continuous, the pre-image $J \subseteq R$ of a closed ideal $J' \subseteq R/I$ is a closed ideal in R. Obviously, J contains I.

Since

$$J = H_I^{-1}(J') = \{j + i \in R \ : \ H_I(j) = [j]_\sim \in J' \ \wedge \ i \in I\},$$

it follows that for a $[j_2]_\sim$ contained within J_2' but not within J_1' the corresponding element $j_2 \in R$ lies in J_2 but not in J_1. Therefore, it results from the inequality $J_1' \neq J_2'$ that $J_1 = H_I^{-1}(J_1') \neq J_2 = H_I^{-1}(J_2')$ as well. Thus, H_I^{-1} is injective.

Step 2: If J is closed, then J' is closed as well.
The image J' of an ideal J containing I is an ideal in the algebra R/I. Furthermore, J is the pre-image of J' because with an element x the whole equivalence class $[x]_\sim$ is contained within J. If the pre-image of an open map is closed, the image is closed as well. Since the homomorphism H_I is open, it follows that J' is closed in R/I. □

The proper ideals of R/I are the images of the proper ideals of R. In particular, the maximal ideals of R/I are the images of the maximal ideals of R.

Theorem 2.16. *The quotient algebra R/M of a commutative normed algebra R with respect to a maximal Ideal M is a field.*

Proof. In order to prove this, it will be shown that there is no proper ideal in R/M which is not trivial (see theorem 2.11 and theorem 2.7). If J were a non-trivial ideal in R/M, its pre-image would be a proper ideal in R which contains M but does not coincide with M. This is a contradiction to the maximality of M. □

The following theorem establishes that the other direction is true as well.

Theorem 2.17. *If the quotient algebra R/I of a commutative algebra R with respect to a proper ideal I is a field, then I is a maximal ideal of R. (The assumption that I is closed is not necessary.)*

Proof. Assuming that J is a proper ideal of R which contains I and does not coincide with I it follows that its image $H_I(J)$ is a non-trivial proper ideal of R/I. Therefore, there exists an element $[x]_\sim$ in the proper ideal $H_I(J)$ which isn't the zero element. Due to theorem 2.7, this is equivalent to the fact that the inverse element of $[x]_\sim$ does not exist. This is a contradiction to the assumption. □

2.4 The canonical isomorphism on the maximal ideals

Definition 2.18 (Spectrum of x)**.** *Let x be a member of a normed algebra R. The set*

$$\sigma(x) = \{\lambda \in \mathbb{C} |\ \nexists (x - \lambda e)^{-1}\}$$

is called the spectrum of the element x.

Theorem 2.19. *Every element x of R has a non-trivial spectrum.*

Proof. Assume that the element $x \in R$ has an empty spectrum, i.e. $(x - \lambda e)^{-1}$ exists for all complex numbers λ. For $\lambda = 0$ the existence of x^{-1} follows.

The element x^{-1} has an empty spectrum as well, i.e. $\exists (x^{-1} - \lambda e)^{-1}\ \forall\ \lambda \in \mathbb{C}$. For $\lambda = 0$, this is obvious. For $\lambda \neq 0$ the existence follows from the expression

$$(x^{-1} - \lambda e)^{-1} = \left[(\frac{1}{\lambda} e - x)(x^{-1}\lambda) \right]^{-1} = -\frac{1}{\lambda} x (x - \frac{1}{\lambda} e)^{-1}.$$

Therefore, $(x^{-1} - \lambda e)^{-1}$ and $(x - \lambda e)^{-1}$ are entire vector functions in λ. Their Taylor series $\sum\limits_{n=0}^{\infty} x^{n+1}\lambda^n$ and $\sum\limits_{n=0}^{\infty} x^{-n-1}\lambda^n$ are absolutely convergent in the complex plane, in particular for $\lambda = 1$. It follows that $\|x^n\|$ and $\|x^{-n}\|$ converge towards 0. This is a contradiction to

$$1 = \|e\| = \|x^n x^{-n}\| \leq \|x^n\| \|x^{-n}\|.$$

□

Theorem 2.20. *A normed field (a normed commutative division algebra) R is isomorphic to the field of complex numbers.*

Proof. Due to the last theorem, there exists a $\lambda \in \mathbb{C}$ for every $x \in R$ such that $(x - \lambda e)$ is not invertible in R. Since R is a field, it follows that $x - \lambda e = 0$. ☐

The application

$$Iso : \begin{cases} R & \rightarrow & \mathbb{C} \\ \lambda e & \mapsto & \lambda \end{cases}$$

defines the isomorphism between the normed field and the complex plane.

Since the quotient algebra with respect to a maximal ideal is a field, the Gelfand-Mazur theorem follows.

Theorem 2.21. *The quotient algebra of a commutative normed algebra R with respect to a maximal ideal M is isomorphic to the field of complex numbers.*

2.4.1 The relation between maximal ideals and complex multiplicative homomorphisms

Theorem 2.22.

- *To every maximal ideal M corresponds a complex multiplicative homomorphism from the algebra R in the field of complex numbers.*
- *To every non-trivial homomorphism from the algebra R in the field of complex numbers corresponds a maximal ideal.*

Proof. First part.

The aforementioned homomorphism is defined by

$$H = (Iso \circ H_M) : \begin{cases} R & \xrightarrow{H_M} & R/M & \xrightarrow{Iso} & \mathbb{C} \\ x & \mapsto & [x]_\sim & \mapsto & \lambda \end{cases}.$$

Second part.

The kernel of a given homomorphism H is an ideal in R. Due to the fundamental theorem on homomorphisms, the factor algebra $R/\ker(H)$ is isomorphic to the image of the homomorphism. Since the image of the non-trivial homomorphism

H is a non-empty subring of the complex number field, the image $H(R)$ must coincide with $\mathbb{C}$. If $H : R \to \mathbb{C}$ is a non-trivial homomorphism, it follows by theorem 2.17 that

- R/M is a field and that
- there exists a maximal ideal M satisfying $\ker(H) = M$.

□

Definition 2.23 (The function $x(M)$ on the maximal ideal space)**.** *Let x be an element of the normed commutative algebra R. The function*

$$x : \begin{cases} \mathcal{M}(R) & \to \quad \mathbb{C} \\ M & \mapsto \quad (Iso \circ H_M)(x) \end{cases}$$

assigns to every maximal ideal M in the space of maximal ideals $\mathcal{M}(R)$ of R the complex number $(Iso \circ H_M)(x)$.

For fixed x and variable maximal ideal M, $x(M)$ is a function on the set of all maximal ideals $\mathcal{M} = \mathcal{M}(R)$ of R. The set of all these applications is denoted with $\hat{R}$.

Properties of the function x(M)

(i). $(x_1 + x_2)(M) = x_1(M) + x_2(M) \quad \forall M \in \mathcal{M}$

(ii). $(x_1 \cdot x_2)(M) = x_1(M) \cdot x_2(M) \quad \forall M \in \mathcal{M}$

(iii). $(\lambda x_1)(M) = \lambda(x_1)(M) \quad \forall M \in \mathcal{M}$

(iv). $e(M) \equiv 1 \quad \forall M \in \mathcal{M}$

(v). x is contained within M if and only if $x(M) = 0$.

(vi). If $M_1 \neq M_2$, there exists an $x \in R$ which separates M_1 and M_2, i.e. $x(M_1) \neq x(M_2)$.

- This separability condition shows that different maximal ideals provide different linear functionals.

(vii). $|x(M)| \leq \|x\| \quad \forall M \in \mathcal{M}$

- $x(M)$ is that number $\lambda_{[x]_\sim}$ which the canonical isomorphism between R/M and the complex number field assigns to the equivalence class $[x]_\sim$ containing x. Since $[x]_\sim = \lambda_{[x]_\sim}[e]_\sim$, it follows that

$$\|[x]_\sim\|_{R/I} = |\lambda_{[x]_\sim}|\|[e]_\sim\|_{R/I} = |\lambda_{[x]_\sim}|$$

  Considering the definition of the norm leads to

$$x(M) = |\lambda_{[x]_\sim}| = \|[x]_\sim\|_{R/I} = \inf_{z\in[x]_\sim} \|z\| \leq \|x\|.$$

The functions $x(M)$ form an algebra $\hat{R}$ with neutral element.

The assertion follows by properties 1 - 4. The application

$$\hat{(\cdot)} : \begin{cases} R & \rightarrow & \hat{R} \\ x & \mapsto & (M \mapsto x(M)) \end{cases}$$

is an homomorphism from R to $\hat{R}$.

$M(x)$ defines a linear bounded multiplicative functional with norm 1.

By properties 1,2,3,4 and 7, $M(x)$ defines for fixed $M \in \mathcal{M}$ and variable $x \in R$ a linear bounded multiplicative functional on R. These linear functionals are to be distinguished from normal linear functionals by their multiplicativity property

$$M(x_1x_2) = M(x_1)M(x_2),$$

which follows from 2.

Due to property 5, the condition for the existence of an inverse element and hence theorem 2.11 can be reformulated.

Theorem 2.24. *An element x of R has an inverse if and only if $x(M)$ does not vanish on $\mathcal{M}(R)$.*

Theorem 2.25. *The spectrum of x coincides with the image of $x(M)$, i.e.*

$$\sigma(x) = \mathrm{Im}(x(M)).$$

Proof. $\supseteq$: If $x(M_0) = \lambda_0$, then $(x - \lambda_0 e)(M_0) = 0$. By property 5, this is equivalent to $(x-\lambda_0 e) \in M_0$. Therefore, $(x-\lambda_0 e)$ has no inverse element.

$\subseteq$: If $(x-\lambda_0 e)^{-1}$ does not exist, $(x-\lambda_0 e)(M)$ vanishes for a maximal ideal M_0, i.e. $x(M_0) = \lambda_0$.

□

2.5 Proof of the univariate case

Theorem 2.26. *Let $f(\zeta)$ be an analytic function on a superset of the spectrum of an element $x \in R$. Let Γ be a rectifiable, simple-closed Jordan-curve, which lies in the area of regularity of $f(\zeta)$ and encloses $\sigma(x)$.*
Under this assumptions the following points hold:

- *The integral $\frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda)}{(\lambda e - x)}d\lambda = \widetilde{f(x)}$ exists and is independent of the choice of Γ.*

- *The canonical homomorphism*

$$\hat{(\cdot)} : \begin{cases} R & \to & \hat{R} \\ x & \mapsto & (M \mapsto x(M)) \end{cases}$$

  maps the element $\widetilde{f(x)}$ of R to the element $f(x(M))$ of $\hat{R}$, which is defined by $\widetilde{f(x)(M)} = f(x(M))$.

- *With $x(M)$, $f(x(M))$ is a member of $\hat{R}$ as well.*

Proof. Since the spectrum of x is enclosed by Γ, the function $z(\lambda) = (\lambda e - x)^{-1}f(\lambda)$ is well defined at any point of Γ. Furthermore, $z(\lambda)$ is continuous with respect to the chosen norm. The integral exists in the sense of norm-convergence and it is independent of the choice of Γ. This is why Cauchy's integral theorem can be applied. For fixed M and variable x, $x(M)$ is a linear functional in x. Therefore,

$$\begin{aligned} \widetilde{f(x)(M)} &= \frac{1}{2\pi i}\int_\Gamma (\lambda e - x)^{-1}(M)f(\lambda)d\lambda \\ &= \frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda)}{(\lambda - x(M))}d\lambda = f(x(M)). \end{aligned}$$

□

2.6 The topology on $\mathcal{M}(R)$

In order to prove the multivariate case, some topological properties of the space $\mathcal{M}(R)$ are required.

In this section, it will be shown that there exists a compact (and unique) Hausdorff topology on the space of maximal ideals $\mathcal{M}(R)$ of an algebra R.

For arbitrarily chosen elements $x_1, x_2, \ldots, x_n$ of the algebra R and for every $\varepsilon > 0$, a neighborhood of the point $M_0 \in \mathcal{M}(R)$ is defined by

$$U_\varepsilon^{x_1,\ldots,x_n}(M_0) = \{M \in \mathcal{M}(R) \mid |x_1(M) - x_1(M_0)| < \epsilon, \ldots, |x_n(M) - x_n(M_0)| < \epsilon\} .$$

These neighborhoods define a fundamental system of neighborhoods which uniquely determines the topology of $\mathcal{M}(R)$.

Definition 2.27 (Fundamental system of neighborhoods $\mathcal{B}(x)$)**.** *Let X be a non-empty set and x an arbitrarily chosen point in X. A non-empty subset $\mathcal{B}(x)$ of $\mathcal{P}(X)$ is called fundamental system of neighborhoods of x, if the following conditions hold.*

(i). $U_1, U_2 \in \mathcal{B}(x) \Rightarrow \exists V \in \mathcal{B}(x) : V \subseteq U_1 \cap U_2$

(ii). $U \in \mathcal{B}(x) \Rightarrow x \in U$

(iii). $U \in \mathcal{B}(x) \Rightarrow \exists V \in \mathcal{B}(x) : \forall y \in V \exists W \in \mathcal{B}(y) : W \subseteq U$

The corresponding topology τ of X is defined as

$$\tau := \{G \in \mathcal{P}(X) \mid \forall x \in G \; \exists U \in \mathcal{B}(x) : U \subseteq G\}$$

Evidently, $\mathcal{B}(M_0) = \{U_\varepsilon^{x_1,\ldots,x_n}(M_0) \mid x_i \in R, \; \forall i \in \{1, \ldots, n\}, \varepsilon > 0\}$ is non-empty and every $U_\varepsilon^{x_1,\ldots,x_n}(M_0)$ contains M_0.

Because of the following conclusion, conditions (1) and (3) are satisfied as well.

- $U_{\min\{\varepsilon_1,\varepsilon_2\}}^{x_1,\ldots,x_n,x_{n+1},\ldots,x_{n+m}}(M_0) \subseteq U_{\varepsilon_1}^{x_1,\ldots,x_n}(M_0) \cap U_{\varepsilon_2}^{x_{n+1},\ldots,x_{n+m}}(M_0)$
- If M_1 is an element of $U_\varepsilon^{x_1,\ldots,x_n}(M_0)$, it follows that $U_\delta^{x_1,\ldots,x_n}(M_1) \subseteq U_\varepsilon^{x_1,\ldots,x_n}(M_0)$, where $0 < \delta < \min\{\varepsilon - |x_1(M_1) - x_1(M_0)|, \ldots, \varepsilon - |x_n(M_1) - x_n(M_0)|\}$.

Since x separates points in $\mathcal{M}(R)$, i.e.

$$M' \neq M \Rightarrow \exists x \in R : x(M') \neq x(M),$$

the topology $\triangle$ of $\mathcal{M}(R)$ is Hausdorff, i.e.

$$U_{\frac{\varepsilon}{2}}^x(M) \cap U_{\frac{\varepsilon}{2}}^x(M') = \emptyset \quad \forall M' : \varepsilon < |x(M') - x(M)|$$

By the definition of the topology $\triangle$ on $\mathcal{M}(R)$, $x(M)$ is continuous, i.e.

$$U_\varepsilon^x(M_0) = \{M \in \mathcal{M}(R) \mid |x(M) - x(M_0)| < \varepsilon\}$$
$$\Longleftrightarrow$$
$$M \in U_\varepsilon^x(M_0) \Rightarrow x(M) \in U_\varepsilon^{\mathbb{C}}(x(M_0)),$$

where $U_\varepsilon^{\mathbb{C}}(x(M_0))$ denotes an open ball with cent re $x(M_0)$ and radius ε in $\mathbb{C}$.

In order to prove that the topology on $\mathcal{M}(R)$ is compact, some topological fundamentals are required. They can be found in [13].

Lemma 2.28. *Let (X, τ) be a topological Hausdorff space and $A \subseteq X$ a compact set. Then, it follows that A is closed.*

Lemma 2.29. *If $f : (X_1, \tau_1) \to (X_2, \tau_2)$ is continuous and $A \subseteq X_1$ compact, it follows that $f(A) \subseteq X_2$ is compact as well.*

Lemma 2.30. *If $f : (X_1, \tau_1) \to (X_2, \tau_2)$ is continuous and bijective, where (X_1, τ_1) is compact and (X_2, τ_2) Hausdorff, it follows that f^{-1} is continuous.*

Lemma 2.31 (Tychonoff)**.** *If $(X_i, \tau_i), i \in I$ is a family of compact topological spaces, it follows that $(\prod_{i\in I} X_i, \prod_{i\in I} \tau_i)$ is a compact topological space as well.*

The main statement of this subsection can now be proved.

Theorem 2.32. *The topological Hausdorff space $(\mathcal{M}(R), \triangle)$ is compact.*

Proof. Step 1: Preparations

A compact disc $Q_x \subseteq \mathbb{C}$ is assigned to every $x \in R$.

- $x \mapsto \overline{U_{\|x\|}^{\mathbb{C}}(0)} =: Q_x$

By Tychonoff's theorem, the Cartesian product Q of all Q_x,

$$Q := \prod_{x\in R} Q_x,$$

is, endowed with the product topology τ, a compact space. The corresponding fundamental system of neighborhoods of a point $(\lambda_x^o)_{x\in R} \in Q$ is given by

$$\left\{(\lambda_x^o)_{x\in R} \in Q \mid |\lambda_{x_1} - \lambda_{x_1}^o| < \varepsilon, \dots, |\lambda_{x_n} - \lambda_{x_n}^o| < \varepsilon\right\}, \quad \varepsilon > 0; x_1, \dots, x_n \in R$$

Step 2: The mapping $f : \begin{cases} \mathcal{M}(R) & \to & Q|_{f(\mathcal{M}(R))} \\ M & \mapsto & (\mu_x)_{x\in R} = (x(M))_{x\in R} \end{cases}$ is well-defined and bijective.

Well-defined: Since $|x(M)| \leq \|x\|$, it follows that

$$\forall M \in \mathcal{M}(R) : \exists (\mu_x)_{x \in R} \in Q, \mu_x = x(M) \quad \forall x \in R$$

Injectivity: For every pair of non-equal maximal ideals exists a $x' \in R$, which assigns different values to those different maximal ideals. Therefore, the corresponding points in Q do not coincide either, at least in the coordinate x', i.e.

$$M_1 \neq M_2 \Rightarrow \exists x'(M_1) \neq x'(M_2)$$
$$\Rightarrow (\mu_x)^1_{x \in R} \neq (\mu_x)^2_{x \in R}, \text{ because } \mu^1_{x'} = x'(M_1) \neq x'(M_2) = \mu^2_{x'}.$$

STEP 3: THE APPLICATION f IS A HOMEOMORPHISM
A comparison between the fundamental systems of neighborhoods of $\mathcal{M}(R)$ and Q shows that f and f^{-1} are continuous.

$\mathcal{M}(\mathbf{R})$: $U^{x_1,\ldots,x_n}_\varepsilon(M_0) = \{M \in \mathcal{M}(R) \mid |x_1(M) - x_1(M_0)| < \epsilon, \ldots, |x_n(M) - x_n(M_0)| < \epsilon\}$

$\mathbf{Q}$: $U^{x_1,\ldots,x_n}_\varepsilon((\lambda^o_x)_{x \in R}) = \{(\lambda^o_x)_{x \in R} \in Q \mid |\lambda_{x_1} - \lambda^o_{x_1}| < \epsilon, \ldots, |\lambda_{x_n} - \lambda^o_{x_n}| < \epsilon\}$

STEP 4: $\mathcal{M}' = f(\mathcal{M}(R))$ IS CLOSED IN THE TOPOLOGY OF Q
It has to be shown that $\mathcal{M}'$ is closed, because then $(\mathcal{M}(R), \triangle)$ is compact and the theorem is proved. If $\mathcal{M}'$ is closed, then it is as closed subset of a compact Hausdorff space compact too. Since f is a homeomorphism, $\mathcal{M}(R) = f^{-1}(f(\mathcal{M}(R)))$ is compact as well.

In order to show that $\mathcal{M}'$ is closed, it will be proved that if $\Lambda = (\lambda^o_x)_{x \in R}$ is an element of the closure $\overline{\mathcal{M}'}$ of $\mathcal{M}'$, Λ belongs to $\mathcal{M}'$ as well, i.e. the equations

$$\lambda^o_{x+y} = \lambda^o_x + \lambda^o_y, c\lambda^o_x = \lambda^o_{cx} \text{ and } \lambda^o_{xy} = \lambda^o_x \lambda^o_y.$$

hold. Since $\Lambda = (\lambda^o_x)_{x \in R}$ is contained within $\overline{\mathcal{M}'}$, there exists a neighborhood

$$V := \left\{(\lambda_x)_{x \in R} \in R \mid |\lambda_e - \lambda^o_e| < \epsilon, |\lambda_x - \lambda^o_x| < \epsilon, |\lambda_y - \lambda^o_y| < \epsilon, |\lambda_{xy} - \lambda^o_{xy}| < \epsilon\right\} = \\ = U^{e,x,y,xy}_\varepsilon(\Lambda)$$

of $\Lambda = (\lambda^o_x)_{x \in R}$ whose intersection with $\mathcal{M}'$ is non-empty and contains an element $(\lambda'_x)_{x \in R}$. Furthermore, there exists a maximal ideal M whose picture is $(\lambda'_x)_{x \in R}$ because the function f is bijective.

$\Lambda \in \overline{\mathcal{M}'} \Rightarrow \exists (\lambda'_x)_{x\in R} \in V \cap \mathcal{M}' \ \wedge \ M \in \mathcal{M}(R) : f(M) = (\lambda'_x)_{x\in R}$

$$|\lambda^o_e - \underbrace{e(M)}_{=\lambda'_e}| = |\lambda^o_e - 1| < \varepsilon$$
$$|\lambda^o_x - \underbrace{x(M)}_{=\lambda'_x}| < \varepsilon, |\lambda^o_y - \underbrace{y(M)}_{=\lambda_y}| < \varepsilon$$
$$|\lambda^o_{xy} - \underbrace{xy(M)}_{=\lambda'_{xy}}| = |\lambda^o_{xy} - x(M)y(M)| < \varepsilon$$

Only the multiplication property, $\lambda^o_{xy} = \lambda^o_x \lambda^o_y$, will be proved.

$$\begin{aligned} |\lambda^o_{xy} - \lambda^o_x\lambda^o_y| &\leq |\lambda^o_{xy} - x(M)y(M) + x(M)y(M) - x(M)\lambda^o_y + \lambda^o_y x(M) - \lambda^o_y\lambda^o_x| \\ &\leq |\lambda^o_{xy} - x(M)y(M)| + \|x\||y(M) - \lambda^o_y| + |\lambda^o_y||x(M) - \lambda^o_x| \\ &< \varepsilon(1 + \|x\| + |\lambda^o_y|) \end{aligned}$$

The other mentioned properties are proved in the same manner.

Therefore, there exists a homomorphism

$$H : \begin{cases} R & \rightarrow & \mathbb{C} \\ x & \mapsto & \lambda^o_x \end{cases}$$

with $ker(H) = M'$, where $x(M') = \lambda_x$ for every $x \in R$. □

2.7 Generating systems and polynomially convex sets

Definition 2.33 (Generating system)**.** *A subset $K \subseteq R$ is a generating system of the algebra R if and only if the smallest closed algebra with neutral element of multiplication which contains K is R.*

- The neutral element is not a generating element.

Theorem 2.34 (Finite number of generating elements)**.** *If the algebra R has a finite number of generating elements, the space $\mathcal{M}(R)$ is homeomorph to a compact set $F \subseteq \mathbb{C}^n$.*

Proof. The set $F = \{z \in \mathbb{C}^n \mid \exists M \in \mathcal{M}(R), z = (x_1(M), \ldots, x_n(M))\}$ is compact because it is the image of a compact set.

The application

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : (\mathcal{M}(R), \triangle) \longrightarrow (F, \|\cdot\|_2)$$

is a continuous and bijective mapping from a compact space into a Hausdorff space and therefore a homeomorphism.

The injectivity is an implication of the following reasoning. If $x_i(M_1) = x_i(M_2)$, $i \in \{1, \ldots, n\}$, all the polynomials in $x_1, \ldots, x_n$ coincide as well, i.e.

$$x_i(M_1) = x_i(M_2), \forall i \Rightarrow \begin{cases} P(x_1(M_1), \ldots, x_n(M_1)) = P(x_2(M_1), \ldots, x_n(M_2)), \\ \text{for all polynomials in } x_1, \ldots, x_n \end{cases}.$$

Since $\{x_1, \ldots, x_n\}$ is a generating system of R, it follows that

$$x(M_1) = x(M_2) \quad \forall x \in R.$$

By the separation property, M_1 must coincide with M_2. □

Definition 2.35 (Polynomially convex set). *A subset F of $\mathbb{C}^n$ is called polynomially convex if and only if for every ζ' in F, there exists a polynomial $P(\zeta_1, \ldots, \zeta_n)$ which satisfies the conditions*

- $P(\zeta') = 1$
- $|P(\zeta)| < 1, \quad \forall \zeta \in F.$

Theorem 2.36. *$F \subseteq \mathbb{C}^n$ is homeomorph to the space of maximal ideals $\mathcal{M}(R)$ of an algebra R with n generating elements if and only if F is closed, bounded and polynomially convex.*

Proof. "⇒": Let R be an algebra generated by n elements z_j, $j \in \{1, \ldots, n\}$. Then $F = \{\zeta \in \mathbb{C}^n \mid \exists M \in \mathcal{M}(R) : \zeta = (x_1(M), \ldots, x_n(M))\}$ is closed and bounded because $\mathcal{M}(R)$ is compact.

Polynomial convexity:

$$\begin{aligned} \zeta' \notin F \quad &\Longleftrightarrow \nexists H = (Iso \circ H_M) : H(x_j) = \zeta'_j \quad \forall j \in \{1, \ldots, n\} \\ &\Longleftrightarrow \nexists M \in \mathcal{M}(R) : (x_j - \zeta'_j e)(M) = 0 \quad \forall j \in \{1, \ldots, n\} \\ &\Longleftrightarrow \nexists M \in \mathcal{M}(R) : x_j - \zeta_j e \in M \quad \forall j \in \{1, \ldots, n\} \end{aligned}$$

Therefore, the ideal

$$\sum_{j=1}^{n}(x_j - \zeta_j' e)r_j, \quad r_j \in R$$

isn't contained within any maximal ideal $M \in \mathcal{M}(R)$ and coincides with the whole algebra.

Since the neutral element of multiplication is a member of every algebra R, it follows that there exist q_j, $j \in \{1, \ldots, n\}$, such that

$$e = \sum_{j=1}^{n}(x_j - \zeta_j' e)q_j.$$

These q_j can be approximated, with arbitrary accuracy, by polynomials in $x_i, i \in \{1, \ldots, n\}$, i.e.

$$\forall \epsilon > 0 \; \exists P_j : \|e - \sum_{j=1}^{n}(x_j - \zeta_j' e)P_j(x_1, \ldots, x_n)\| < \varepsilon.$$

The corresponding polynomial in $(\zeta_1, \ldots, \zeta_n)$, defining the polynomially convex set, is

$$1 - \sum_{j=1}^{n}(\zeta_j - \zeta_j' e)P_j(\zeta_1, \ldots, \zeta_n).$$

"$\Leftarrow$": It will be shown that a closed, bounded and polynomially convex subset F of $\mathbb{C}^n$ is homeomorph to a space of maximal ideals of an algebra R.

Let R' be the algebra of all polynomials $P(\zeta_1, \ldots, \zeta_n)$, which is endowed with the norm $\|P\| = \max_{\zeta \in F}|P(\zeta_1, \ldots, \zeta_n)|$. R is the completed version of R' (with respect to $\|\cdot\|$) and the $z_j(\zeta) = \zeta_j$, $j \in \{1, \ldots, n\}$, are the generating elements of R.

Since the algebra R has n generating elements, a compact set F', which is homeomorph to $\mathcal{M}(R)$, exists. Due to the following reasoning, the set F coincides with F'.

$F \subseteq F'$: Every $\zeta^0 \in F$ defines by the homomorphism $z_j(\zeta) \mapsto \zeta_j^0$ a maximal ideal in R. Therefore, ζ_0 is contained within $F' \cong \mathcal{M}(R)$ as well.

Injectivity: Different points in F give different maximal ideals.

$$\zeta^1 \neq \zeta^2 \Rightarrow \exists j \in \{1, \ldots, n\} : H_j^1(z_j) = \zeta_j^1 \neq \zeta_j^2 = H_j^2(z_j) \Rightarrow M_{\zeta^1} \neq M_{\zeta^2}$$

$F \supseteq F'$: If ζ' is not contained within F, there exists a polynomial whose absolute value is smaller than one for every ζ in F and whose value at ζ' is one. Since F is compact, the norm of the polynomial is smaller than one. Therefore, a maximal ideal which corresponds to ζ' does not exist. In mathematical terms, this means

$\zeta' \notin F \Rightarrow \exists P(\zeta_1, \ldots, \zeta_n) : P(\zeta) < 1 \; \forall \zeta \in F, \; \wedge \; P(\zeta') = 1$

$\Rightarrow \|P\| = \max_{\zeta \in F} |P(\zeta)| = c < 1$

$\Rightarrow |P(z_1, \ldots, z_n)(M)| \leq \|P(z_1, \ldots, z_n)\| = c < 1 \Rightarrow \zeta' \notin \mathcal{M}(R) \cong F'$.

□

2.8 Proof of the multivariate case

The purpose of this section is the proof of the following theorem.

Theorem 2.37. *If $f(\zeta_1, \ldots, \zeta_n)$ is analytic on the set*

$$\sigma_R(x_1, \ldots, x_n) = \{\zeta \in \mathbb{C}^n \mid \exists M \in \mathbb{M}(R) : \zeta = (x_1(M), \ldots, x_n(M))\},$$

then there exists an $x \in R$ satisfying

$$f(x_1(M), \ldots, x_n(M)) = x(M), \quad \forall M \in \mathcal{M}(R).$$

In a first step, it will be proved for the case where the algebra is generated by a finite number of elements. Subsequently, this result is generalized to an arbitrary normed commutative algebra.

2.8.1 Case A: Algebra R is generated by a finite number of elements

In analogy to the proof of the one dimensional case, the analytically transformed element will be represented as a multidimensional integral on a certain domain. The domain contains the compact and polynomially convex set, which is isomorphic to the space of maximal ideals. Since the algebra is complete, this integral is a member of the algebra as well. The integral representation will be proved by Weil's theorem 2.39.

Definition 2.38 (Weil domain)**.** *The finite intersection of sets of the form $G_\nu = \{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta_1, \ldots, \zeta_n)| < 1\}$, i.e. the set $G = \bigcap_{\nu=1}^{N} G_\nu$, is a Weil domain if and only if the real dimension of $\bigcap_{k=1}^{n} \{\zeta \in \mathbb{C}^n \mid |P_{i_k}(\zeta_1, \ldots, \zeta_n)| = 1\}$ is smaller than n for all $\{i_1, \ldots, i_n\} \subseteq \{1, \ldots, N\}$.*

Weil's theorem

The Weil domain is defined by inequalities of the form

$$|P_i(\zeta_1, \ldots, \zeta_n)| < 1, \quad i \in \{1, \ldots, N\}.$$

The intersection of the boundary of n such domains, $\bigcap_{j=1}^{n} \left\{\zeta \in \mathbb{C}^n \mid |P_{i_j}(\zeta)| = 1\right\}$, is denoted by $\sigma_{i_1 \cdots i_n}$. This intersection is endowed with a certain orientation.

For every point $(\tau_1, \ldots, \tau_n) \in F$, the difference of two polynomials can be represented as

$$P_i(\zeta_1, \ldots, \zeta_n) - P_i(\tau_1, \ldots, \tau_n) = \sum_{j=1}^{n} (\zeta_j - \tau_j) Q_{ij}(\zeta_1, \ldots, \zeta_n),$$

where the Q_{ij} are polynomials in $(\zeta_1, \ldots, \zeta_n)$ and its coefficients depend on $(\tau_1, \ldots, \tau_n)$.

Theorem 2.39 (Weil's integral representation)**.** *Let $f(\zeta_1, \ldots, \zeta_n)$ be an analytic function on a domain $G \subseteq \mathbb{C}^n$ which contains a Weil domain $\mathcal{G}$. Furthermore, $\tau = (\tau_1, \ldots, \tau_n)$ is an element of $F \subseteq \mathcal{G}$ and $D_{i_1 \cdots i_n}$ denotes* $\det \left[(Q_{i_\alpha j})_{\alpha,j=1}^{n} \right]$. *Then, this function can be represented as*

$$f(\tau_1, \ldots, \tau_n) =$$

$$\sum_{\substack{i_1, \ldots, i_n = 1 \\ i_1 < \cdots < i_n}}^{N} \left(\frac{1}{2\pi i} \right)^n \int_{\sigma_{i_1 \cdots i_n}} \frac{D_{i_1 \cdots i_n} f(\zeta_1, \ldots, \zeta_n)}{\prod_{j=1}^{n} P_{i_j}(\zeta_1, \ldots, \zeta_n) - P_{i_j}(\tau_1, \ldots, \tau_n)} d\zeta_1 \cdots d\zeta_n.$$

Proof. See B.A. Fuks [5]. □

Interpretation in the algebra

- Since $P_{i_j}(\zeta_1, \ldots, \zeta_n)e - P_{i_j}(x_1, \ldots, x_n)$ does not vanish for $\zeta \in \sigma_{i_1 \cdots i_n}$ and τ is contained within F, it follows that $\left[P_{i_j}(\zeta_1, \ldots, \zeta_n)e - P_{i_j}(x_1, \ldots, x_n) \right]^{-1}$ exists as continuous function in ζ.
- Since $D_{i_1 \cdots i_n}$ is a polynomial in ζ for fixed τ or fixed elements $x_1, \ldots, x_n \in R$, $D_{i_1 \cdots i_n}(\zeta) f(\zeta) \left[\prod_{j=1}^{n} P_{i_j}(\zeta_1, \ldots, \zeta_n)e - P_{i_j}(x_1, \ldots, x_n) \right]^{-1}$ can be integrated over $\sigma_{i_1 \cdots i_n}$ with respect to ζ.

The integral is a member of R as well, because the algebra is complete. Since the canonical homomorphism H_M is continuous, the values of the obtained element in R coincide on $\mathcal{M} \cong F$ with the values of $f(\tau_1, \ldots, \tau_n) = f(x_1(M), \ldots, f(x_n(M)))$ in the corresponding points of $F \cong \mathcal{M}(R)$.

In order to apply Weil's theorem, the Weil domain will now be constructed.

Construction of the Weil domain.

$\mathcal{M}(R)$ is homeomorph to a compact and polynomially convex subset F of $\mathbb{C}^n$. The function $f(\zeta)$ is analytic on an open superset G of F which can be written as intersection of sets of the form

$$G_\nu = \{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta_1, \ldots, \zeta_n)| < 1\},$$

because it is polynomially convex. The Weil domain is the intersection of a finite number of sets from $F = \bigcap_{\nu=1}^{\infty} G_\nu$. These sets will be chosen such that the Weil domain is contained within the area of regularity of $f(\zeta)$ and that the real dimension of its boundary is smaller than n for all possible choices of n indexes.

The Weil domain $\mathcal{G}$ will be constructed by a compactness argument. Since F is compact, the maxima m_ν of the polynomials P_ν exist and they are smaller than 1. Hence, F can be represented as intersection of closed, even compact sets F_ν by choosing a constant ϑ_ν bigger than the maximum m_ν of the polynomial $P_\nu(\zeta)$ and smaller than 1.

- $\mathcal{M}(R) \cong F \subseteq F_\nu \subseteq G_\nu$
  - $\max_{\zeta \in F} |P_\nu(\zeta_1, \ldots, \zeta_n)| = m_\nu < \vartheta_\nu < 1$
  - $F_\nu = \{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta_1, \ldots, \zeta_n)| \le \vartheta_\nu\}$
  - $G_\nu = \{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta_1, \ldots, \zeta_n)| < 1\}$
    - Since $F = \bigcap_\nu G_\nu$ and $F_\nu \subseteq G_\nu$ it follows that $F = \bigcap_\nu F_\nu$.

In the same manner, the generating elements assume their maxima. Therefore, the sets Q_j and S_j, created from the generating elements, can be defined.

- $\mathcal{M}(R) \cong F \subseteq S_j \subseteq Q_j$
  - $\max_{M \in \mathcal{M}(R)} |z_j(M)| = \max_{\zeta \in F} |\zeta_j| < a_j < c_j$
  - $S_j = \{\zeta \in \mathbb{C}^n \mid |\zeta_j| \le a_j\}$
  - $Q_j = \{\zeta \in \mathbb{C}^n \mid |\zeta_j| < c_j\}$

Since the complement G^c of G and the sets F_ν are closed in $\mathbb{C}^n$, they are closed in the relative topology of the compact set $S = \bigcap_{j=1}^{n} S_j$ as well. If the intersection of a family of sets in a compact space is empty, a finite number of sets can be chosen, whose intersection is empty too, i.e.

$$\bigcap_{\nu} F_\nu \cap G^c \cap S = \emptyset \Longrightarrow \exists \nu_1, \dots, \nu_n : \bigcap_{i=1}^{m} F_{\nu_i} \cap G^c \cap S = \emptyset$$
$$\Longrightarrow \bigcap_{i=1}^{m} F_{\nu_i} \cap S \subseteq G.$$

Therefore, the Weil domain $\mathcal{G}$ in demand is

$$\mathcal{G} = \bigcap_{i=1}^{m} F_{\nu_i}^o \cap \bigcap_{j=1}^{n} S_j^o,$$

where A^o denotes the interior of A. In order to prove that $\mathcal{G}$ is a Weil domain, the second condition in its definition must be verified. It must be shown that the real dimension of the boundary is smaller than n.

The real dimension of the boundary is smaller than n, i.e. the P_ν can be chosen not being functional dependent:

The polynomial $\widetilde{P_\nu(\zeta)}$ is defined by

$$\widetilde{P_\nu(\zeta)} = \frac{1}{\vartheta_\nu} P_\nu(\zeta) + R_\nu(\zeta),$$

where the absolute value of $R_\nu(\zeta)$ is not bigger than ε on

$$Q = \bigcap_{j=1}^{n} Q_j = \{\zeta \in \mathbb{C}^n \mid |\zeta_j| < c_j, j \in \{1, \dots, n\}\} \supseteq F.$$

It will be shown that an $\varepsilon > 0$ and polynomials $\widetilde{P_\nu(\zeta)}$ exists, such that the intersection of Q and a set generated by $\widetilde{P_\nu(\zeta)}$ is contained within the set generated by $P_\nu(\zeta)$, i.e.

$$F \subseteq \left(\left\{ \zeta \in \mathbb{C}^n \mid |\widetilde{P_\nu(\zeta)}| < 1 \right\} \cap Q \right) \subseteq \{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta)| < 1\}, \quad \forall \, \nu.$$

First Inclusion: Evidently, an $\varepsilon > 0$ can be chosen such that the norm of $\widetilde{P_\nu(\zeta)}$ is smaller than one.

$$\max_{\zeta \in F} |\widetilde{P_\nu(\zeta)}| \le \frac{1}{\vartheta_\nu} \max_{\zeta \in F} |P_\nu(\zeta)| + \max_{\zeta \in F} |\widetilde{R_\nu(\zeta)}| \le \underbrace{\frac{m_\nu}{\vartheta_\nu}}_{<1} + \varepsilon$$

Hence, if ζ is an element of F, it is an element of $\left\{ \zeta \in \mathbb{C}^n \mid |\widetilde{P_\nu(\zeta)}| < 1 \right\} \cap Q$ as well.

Second Inclusion: If the norm of $\widetilde{P_\nu(\zeta)}$ is smaller than 1, it follows that an $\varepsilon > 0$ exists such that $P_\nu(\zeta)$ is smaller than one as well.

$$\|P_\nu(\zeta)\| = \|\vartheta_\nu(\widetilde{P_\nu(\zeta)} - R_\nu(\zeta))\| < \vartheta_\nu(1+\varepsilon)$$

Hence, if ζ is an element of $\left\{\zeta \in \mathbb{C}^n \mid |\widetilde{P_\nu(\zeta)}| < 1\right\} \cap Q$, it is contained within $\{\zeta \in \mathbb{C}^n \mid |P_\nu(\zeta)| < 1\}$ as well.

Therefore, one can chose the polynomials $\widetilde{P_\nu(\zeta)}$ in a way that they are not functionally dependent on each other and that the real dimension of

$$\bigcap_{k=1}^{n} \left\{\zeta \in \mathbb{C}^n \mid |P_{i_k}(\widetilde{\zeta_1, \dots}, \zeta_n)| = 1\right\}, \quad \{i_1, \dots, i_n\} \subseteq \{1, \dots, N\}$$

is smaller than n.

2.8.2 General Case

Definition 2.40 (Common spectrum)**.** *The set*

$$\sigma_R(x_1, \dots, x_n) = \{\zeta \in \mathbb{C}^n \mid \exists M \in \mathcal{M}(R), \zeta = (x_1(M), \dots, x_n(M))\}$$

is called the common spectrum of n elements $x_1, \dots, x_n$ of R..

- $\sigma_R(x_1, \dots, x_n)$ is the image of a compact set and therefore compact as well. Moreover, the common spectrum is contained within $\{\zeta \in \mathbb{C}^n \mid |\zeta_j| \leq \|x_j\|, j \in \{1, \dots, n\}\}$.
- When $R_0 \subseteq R$ is the subalgebra generated from $x_1, \dots, x_n$, in general, only the inclusion

$$\sigma_R(x_1, \dots, x_n) \subseteq \sigma_{R_0}(x_1, \dots, x_n)$$

  holds. This is because the intersection of the maximal ideals of R with R_0 is a maximal ideal in R_0.
- If the elements $x_1, \dots, x_n$ generate R, $\sigma_R(x_1, \dots, x_n)$ coincides with $\mathcal{M}(R) \cong F$. Thus, theorem 2.37 follows directly from Case A.

For every analytic function $f(\zeta)$ on $\sigma_R(x_1, \dots, x_n)$ an open set $U \supseteq \sigma_R(x_1, \dots, x_n)$ exists, where $f(\zeta)$ is analytic as well.

It will be proved that a subalgebra $R' \subseteq R$ exists which is generated by finitely many elements $x_1, \ldots, x_n, y_1, \ldots, y_p$ such that the projection $\sigma_{R'}(x_1, \ldots, x_n) \supseteq \sigma_R(x_1, \ldots, x_n)$ of the common spectrum $\sigma_{R'}(x_1, \ldots, x_n y_1, \ldots y_p) \subseteq \mathbb{C}^{n+p}$ on $\mathbb{C}^n$ is contained within the open set U as well. Firstly, it will be shown that for every point, which isn't contained within $\sigma_R(x_1, \ldots, x_n)$, a whole neighborhood of this point isn't contained within $\sigma_{R_1}(x_1, \ldots, x_n)$, where $R_1 \subseteq R' \subseteq R$. Secondly, a compactness argument will be used to construct R' such that $\sigma_{R'}(x_1, \ldots, x_n) \subseteq U$.

Then, $f(\zeta)$ can be continued constantly across $\mathcal{M}(R') \supseteq \mathcal{M}(R)$ on the remaining p arguments. Due to case A, there exists an x in R' with the property in demand.

Step 1: Choice of $y_1, \ldots, y_p$.

The circumstance that ζ' is not an element of $\sigma_R(x_1, \ldots, x_n)$ is equivalent to the fact that there is no maximal ideal in R, which contains all differences $x_j - \zeta'_j e$, $j \in \{1, \ldots, n\}$, i.e.

$$\begin{aligned} \zeta' \notin \sigma_R(x_1, \ldots, x_n) \quad &\Longleftrightarrow \nexists H = (Iso \circ H_M) : H(x_j) = \zeta'_j \quad \forall j \in \{1, \ldots, n\} \\ &\Longleftrightarrow \nexists M \in \mathcal{M}(R) : (x_j - \zeta'_j e)(M) = 0 \quad \forall j \in \{1, \ldots, n\} \\ &\Longleftrightarrow \nexists M \in \mathcal{M}(R) : x_j - \zeta'_j e \in M \quad \forall j \in \{1, \ldots, n\}. \end{aligned}$$

Therefore, no maximal ideal contains the ideal

$$\sum_{j=1}^{n} (x_j - \zeta'_j e) q_j, \quad q_j \in R,$$

generated by $\{(x_j - \zeta'_j e) \mid j \in \{1, \ldots, n\}\}$. Hence, this ideal coincides with the whole algebra.

Furthermore, there exist elements $y_1, \ldots, y_n \in R$ such that

$$e = \sum_{j=1}^{n} (x_j - \zeta'_j e) y_j.$$

In a neighborhood $U_\varepsilon^{\mathbb{C}^n}(\zeta')$ the difference between the neutral element and $\sum_{j=1}^{n} (x_j - \zeta_j e) y_j$ is smaller than 1 for every $\zeta \in U_\varepsilon^{\mathbb{C}^n}(\zeta')$. For this reason, it has an inverse element, i.e.

$$\exists U_\varepsilon^{\mathbb{C}^n}(\zeta') : \|e - \sum_{j=1}^{n} (x_j - \zeta_j e) y_j\| < 1, \quad \forall \zeta \in U_\varepsilon^{\mathbb{C}^n}(\zeta').$$

Now, since $e, x_1, \ldots, x_n, y_1, \ldots, y_p$ and their powers are already contained within the subalgebra $R_1 \subseteq R$, generated by these elements, the inverse element is even contained within R_1. By theorem 2.11, it follows that $(x - \zeta e)$ isn't a member of any maximal ideal M of $\mathcal{M}(R_1)$. This is equivalent to the fact that $x(M) \neq \zeta$ for all maximal ideals M in $\mathcal{M}(R_1)$ and that ζ isn't contained within $\sigma_{R_1}(x_1, \ldots, x_n)$.

Therefore,

$$\zeta \notin \sigma_{R_1}(x_1, \ldots, x_n), \quad \forall \zeta \in U_\varepsilon^{\mathbb{C}^n}(\zeta')$$

and hence

$$\zeta \notin \sigma_{R'}(x_1, \ldots, x_n), \quad \forall \zeta \in U_\varepsilon^{\mathbb{C}^n}(\zeta'), \forall R' \supseteq R_1.$$

Step 2: Compactness argument.

The set theoretic difference $Z \backslash U$ between

$$Z = \{\zeta \in \mathbb{C}^n \mid |\zeta_j| \leq \|x_j\|, j \in \{1, \ldots, n\}\}$$

and U is compact and can be covered by finitely many of the above constructed neighborhoods $U_{\varepsilon_i}^{\mathbb{C}^n}(\zeta'^{(i)})$.

R' is the algebra generated by the elements $x_1, \ldots, x_n, y_1, \ldots, y_n, \ldots, y_p$ which belong to the finite number of the open balls $U_{\varepsilon_i}^{\mathbb{C}^n}(\zeta'^{(i)})$ in the finite cover of $Z \backslash U$. Due to the above reasoning, the projection of the common spectrum of $x_1, \ldots, x_n, y_1, \ldots, y_n, \ldots, y_p$ on the space of the first n components, i.e. $\sigma_{R'}(x_1, \ldots, x_n)$ does not contain any point of $Z \backslash U$ and is therefore contained within U. □

3 Perturbation of Polynomials

In this chapter, the behavior of polynomials under perturbation of their coefficient vector will be investigated. In particular, the question whether the roots of the characteristic polynomial depend continuously on the coefficients will be answered positively. The whole theory is developed in the frame of complex analysis. Firstly, the structure of critical points will be investigated. Critical points are those points for which the number of distinct roots decreases. Secondly, the notion of cycles of roots will be introduced. The cycles, which evolve from analytic continuation of roots, will play a crucial role in the behavior of the corresponding eigenprojections.

Furthermore, it has been shown in Kato [11], that the eigenvalues can be represented as continuous functions of a real parameter if the coefficients of the (complex) matrix depend continuously on it. In the case of a complex parameter, the representation as continuous or analytic function of that complex parameter is in general not possible if multiple roots occur.

Definition 3.1 (Polynomial of order n)**.** *The function*

$$p(s,a) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0 \quad \in \mathbb{C}[s] \tag{6}$$

is called a polynomial of order n with coefficient vector $a = (a_0, \ldots, a_n) \in \mathbb{C}^{n+1}$. Furthermore, $n(a) := \deg(p(s,a))$ denotes the degree of the polynomial, i.e. $a_{n(a)} \neq 0$ and $a_j = 0$ for $j = n(a)+1, \ldots, n$. The coefficient $a_{n(a)}$ is called the leading coefficient.

If the whole unordered n-tuple, i.e. a set in which the multiplicities are taken into account, of the polynomial $p(s,a)$ is considered, continuous dependence of these roots on the coefficient vector with respect to the norm defined by (7) can be established.

The unordered n-tuple of roots of $p(s,a)$ is for any $a = (a_n, \ldots, a_0) \in \mathbb{C}^{n+1}$ with $a_n \neq 0$ denoted by

$$\Lambda(a) = \lfloor s_1(a), \ldots, s_n(a) \rfloor.$$

Furthermore, if $A \in \mathbb{C}^{n\times n}$ is a matrix, then $\Lambda(A)$ denotes the unordered n-tuple of eigenvalues of A, i.e. the roots of the characteristic polynomial $\chi_A(s)$, where multiplicities are taken into account.

Definition 3.2 (Distance between unordered n-tuples)**.** *A distance between unordered n-tuples is defined by*

$$d(\Lambda, \Lambda') = \min_{\pi \in \Pi_n} \max_{k \in \{1,\dots,n\}} |\lambda_{\pi(k)} - \lambda'_k|, \tag{7}$$

where Λ and Λ' are unordered n-tuples comprising respectively λ and λ' and Π_n the group of permutations of the set $\{1,\dots,n\}$. This metric is called matching distance.

The corresponding metric space of all unordered complex n-tuples endowed with the metric (7) is $(\mathcal{T}_n(\mathbb{C}), d)$.

The following theorem claims that a small perturbation of the coefficient vector a causes only a small perturbation of the roots and that, if additional roots occur, their distance from zero is large.

Theorem 3.3. *Let $p(s, \tilde{a})$ be a non-constant polynomial whose degree $n(\tilde{a})$ is smaller than or equal to n and which has l distinct roots $\tilde{s}_j$, $j \in \{1,\dots,l\}$, with corresponding multiplicities m_j, $j \in \{1,\dots,l\}$.*

Then, for the mutually disjoint closed disks $\overline{U_\varepsilon(\tilde{s}_j)}$ a $\delta(\varepsilon) > 0$ exists, such that for all coefficient vectors a in the open ball $U_{\delta(\varepsilon)}(\tilde{a})$ there are exactly m_j roots of $p(s, a)$ inside the disk $U_\varepsilon(\tilde{s}_j)$ for $j \in \{1,\dots,l\}$ and $n(a) - n(\tilde{a})$ roots outside $U_{\frac{1}{\varepsilon}}(0)$.

Proof. STEP 1: THE m_j ROOTS ASSOCIATED TO $\tilde{s}_j$ REMAIN IN $U_\varepsilon(\tilde{s}_j)$.
For the proof Rouché's theorem will be used:

Theorem (Rouché)**.** *Let f and g be meromorphic functions on an open and connected subset G of $\mathbb{C}$. The closed disk $\overline{U_r(a)}$ with cent re a is contained within G. Furthermore, f and g do not have poles or zeros on the boundary $\partial U_r(a)$ of $U_r(a)$. It follows from the inequality*

$$|f(z) - g(z)| < |f(z)| + |g(z)| \quad \forall z \in \partial U_r(a)$$

that the difference between zeros and poles of the function f equals the the difference between zeros and poles of g on the whole open disk $U_r(a)$, i.e.

$$N_f - P_f = N_g - P_g,$$

where N_f and P_f denote the number of zeros and poles of f in $U_r(a)$.

Proof. See [3]. □

Choose an $\varepsilon > 0$ such that the disks with the distinct zeros as centers are mutually disjoint, i.e. $\overline{U_\varepsilon(\tilde{s}_j)} \cap \overline{U_\varepsilon(\tilde{s}_i)} = \emptyset \quad \forall i \neq j$, and define

$$\mu_j(\varepsilon) = \min_{s \in \partial U_\varepsilon(\tilde{s}_j)} |p(s, \tilde{a})| > 0.$$

By continuous dependence of the polynomial on the coefficient vector a, a disk with cent re $\tilde{a}$ and radius $\delta(\mu_j(\varepsilon)) > 0$ can be chosen such that every polynomial $p(s, a)$, $a \in U_{\delta(\mu_j(\varepsilon))}(\tilde{a})$, differs only little from the initial polynomial, i.e.

$$\sup_{s \in \partial U_\varepsilon(\tilde{s}_j)} |p(s, a) - p(s, \tilde{a})| < \mu_j(\varepsilon) \qquad \forall j \in \{1, \dots, l\}, \quad \forall a \in U_{\delta(\mu_j(\varepsilon))}(\tilde{a})$$

It follows with Rouché's theorem that the number of zeros of the two polynomials $p(s, a)$ and $p(s, \tilde{a})$ coincide on each $U_\varepsilon(\tilde{s}_j)$ and for every $a \in U_{\delta(\mu_j(\varepsilon))}(\tilde{a})$, i.e.

$$N_{p(s,a)} = N_{p(s,\tilde{a})} \text{ in } U_\varepsilon(\tilde{s}_j), \forall a \in U_{\delta(\mu_j(\varepsilon))}(\tilde{a}).$$

Step 2: $n(a) - n(\tilde{a})$ roots are outside of $U_{\frac{1}{\varepsilon}}(0)$.
For the second part it is assumed that

- a sequence $(a_k)_{k \in \mathbb{N}}$ in $U_\delta(\tilde{a})$ exists which converges versus $\tilde{a}$, i.e. $a_k \xrightarrow{k \to \infty} \tilde{a}$, and that
- for every a_k, there exists a root z_k of $p(\cdot, a_k) \in U_{\frac{1}{\varepsilon}}(0) \backslash \bigcup_{j=1}^{l} U_\varepsilon(\tilde{s}_j)$.

Since the closure of $U_{\frac{1}{\varepsilon}}(0) \backslash \bigcup_{j=1}^{l} U_\varepsilon(\tilde{s}_j)$ is compact, one can choose a subsequence $(z_k)_{k \in \mathbb{N}}$ which is convergent in $\mathbb{C}$. But then, the limit $z^0 \in \mathbb{C} \backslash \bigcup_{j=1}^{l} U_\varepsilon(\tilde{s}_j)$ would satisfy $p(z^0, \tilde{a}) = \lim_{k \to \infty} p(z_k, a_k) = 0$ and hence $p(\cdot, \tilde{a})$ would have more than $n(\tilde{a})$ roots. This contradiction shows that the second assertion also holds. □

By the following corollary, the meaning of the expression "the roots depend continuously on the coefficient vector" is specified.

Corollary 3.4 (Roots depend continuously on coefficient vector)**.** *The map*

$$\Lambda : \begin{cases} \{x \in \mathbb{C}^{n+1} | x_n \neq 0\} & \to & (\mathcal{T}_n(\mathbb{C}), d) \\ a & \mapsto & \Lambda(a) \end{cases}$$

associates to every coefficient vector $a \in \left\{x \in \mathbb{C}^{n+1} \mid x_n \neq 0\right\}$ *the unordered n-tuple of the roots of* $p(s, a)$*. This map is continuous.*

Proof. This corollary follows directly from theorem 3.3. ☐

Since $\mathbb{C}^{n \times n}$ is isomorphic to the space of all linear applications from $\mathbb{C}^n$ to $\mathbb{C}^n$, which is denoted by $\mathcal{L}(\mathbb{C}^n)$, a linear operator $A \in \mathcal{L}(\mathbb{C}^n)$ is represented as a matrix in respect to the standard basis of $\mathbb{C}^n$. This space is endowed with an arbitrarily chosen operator norm $\|\cdot\| = \|\cdot\|_{\mathcal{L}(\mathbb{C}^n)}$.

Corollary 3.5 (Results for matrices)**.** *The map*

$$\Lambda : \begin{cases} \mathbb{C}^{n \times n} & \rightarrow \quad (\mathcal{T}_n(\mathbb{C}), d) \\ A & \mapsto \quad \Lambda(A) \end{cases}$$

assigns to every $A \in \mathbb{C}^{n \times n}$ the unordered n-tuple of roots of its characteristic polynomial $\xi_A(s) = det(sI - A)$. This map is continuous on $\mathbb{C}^{n \times n}$, i.e. for any given $A_0 \in \mathbb{C}^{n \times n}$ and $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\|A - A_0\| < \delta \Rightarrow \operatorname{dist}(\Lambda(A), \Lambda(A_0)) < \varepsilon,$$

where the distance between unordered n-tuples is defined by (7).

Proof. This corollary follows directly from theorem 3.3 applied on the characteristic polynomial $\xi_A(s)$ of A. ☐

This corollary does not mean that the eigenvalues of $A \in \mathbb{C}^{n \times n}$ can be represented as continuous functions of A.

The following theorem shows that the roots can be represented as analytic functions if they are simple.

Theorem 3.6 (Simple root representation)**.** *If $\tilde{s}_j$ is a root of multiplicity one of $p(s, \tilde{a})$, then there exist $\delta, \varepsilon > 0$ such that for all coefficients vectors a in $U_\delta(\tilde{a})$ the following statements hold.*

- *The polynomial $p(s, a)$ has exactly one root $s_j(a)$ in $U_\varepsilon(\tilde{s}_j)$.*
- *The function $s_j(a)$ depends analytically on a in $U_\delta(\tilde{a})$.*
- *The function $s_j(a)$ coincides with $\tilde{s}_j$ on the point $\tilde{a}$, i.e. $s_j(\tilde{a}) = \tilde{s}_j$.*

Proof. This theorem follows from theorem 3.3 and from the residue theorem, by which

$$\frac{1}{2\pi i}\int_{\Gamma_j}\frac{sp'(s,a)}{p(s,a)}ds = s_j(a), \qquad a \in U_\delta(\tilde{a})$$

is analytic with respect to a.The prime in the expression $p'(s,a)$ denotes $\frac{d}{ds}p(s,a)$. ☐

The same result can be stated for matrices and their characteristic polynomials.

Corollary 3.7. *If $\lambda_0 \in \mathbb{C}$ is a simple eigenvalue of $A_0 \in \mathbb{C}^{n\times n}$, then there exist $\delta, \varepsilon > 0$ such that for all matrices A in $U_\delta(A)$, the following statements hold.*

- *All $A \in U_\delta(A_0)$ have exactly one simple eigenvalue $\lambda(A) \in U_\varepsilon(\lambda_0)$.*
- *The eigenvalue function $\lambda(A)$ depends analytically on the entries of A in $U_\delta(A_0)$.*
- *The function $\lambda(A)$ coincides in A_0 with λ_0, $\lambda(A_0) = \lambda_0$*

Proof. This corollary follows directly from theorem 3.6 applied to the characteristic polynomial $\xi_A(s)$. ☐

3.1 Polynomials with Meromorphic Coefficients

In this subsection, the coefficients of a polynomial are members of the field of meromorphic functions $\mathcal{M}(\Omega)$ or members of the ring of holomorphic functions $\mathcal{O}(\Omega)$ defined on Ω.

In the following, Ω denotes a domain, i.e. a connected and open set, in the field of complex numbers. Due to the identity theorem, the ring $\mathcal{O}(\Omega)$ of analytic functions on Ω is even an integral domain, i.e. a ring with multiplicative unit and without zero divisors. Therefore, $\mathcal{M}\Omega$ is the quotient field of $\mathcal{O}(\Omega)$.

Furthermore, $\mathcal{O}(\Omega)[s]$ and $\mathcal{M}(\Omega)[s]$ are respectively the rings of all the polynomials with coefficients in $\mathcal{O}(\Omega)$ or $\mathcal{M}(\Omega)$.

In order to prove the theorems about the polynomials with coefficients in an arbitrary field, it is necessary to know some facts about these polynomials.

Definition 3.8 (Irreducibility)**.** *The polynomial* $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ *is called irreducible over* $\mathcal{M}(\Omega)$ *if and only if* $p(s, a(z))$ *is non-constant and cannot be represented as a product of two non-constant polynomials, i.e.*

- $\deg(p) \geq 1$
- $\nexists q \in \mathcal{M}(\Omega) : 0 < \deg(q) < \deg(p) \ \wedge \ q|p$, *where* $deg(p)$ *denotes the degree of the polynomial* $p(s, a(z))$.

Definition 3.9 ($p(s, a(z))$ is prime)**.** $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ *is prime over* $\mathcal{M}(\Omega)$ *if and only if it follows from the fact that* $p(s, a(z))$ *divides the product* $q_1(s, b(z)) \cdot q_2(s, c(z))$ *in* $\mathcal{M}(\Omega)[s]$ *that* $p(s, a(z))$ *divides* $q_1(s, b(z))$ *or* $q_2(s, c(z))$.

Definition 3.10 ($p(s, a(z))$ and $p(s, b(z))$ are coprime)**.** *Two polynomials* $p(s, a(z))$, $p(s, b(z)) \in \mathcal{M}(\Omega)[s]$ *are called coprime over* $\mathcal{M}(\Omega)$ *if and only if their greatest common divisor is a non-zero constant polynomial.*

Theorem 3.11. *If* $p \in \mathcal{M}(\Omega)[s]$ *is irreducible, then* p *is prime.*

Definition 3.12 (Resultant matrix of two polynomials)**.** *The resultant matrix* $R(p, q) \in \mathcal{M}(\Omega)^{(n+m)\times(n+m)}$ *of the polynomials* $p(s, a(z)) = \sum_{k=0}^{n} a_k(z)s^k$ *and* $q(s, b(z)) = \sum_{k=0}^{m} b_k(z)s^k$ *is defined as*

$$R(p,q) = \begin{pmatrix} a_n(z) & \cdots & a_0(z) & 0 & \cdots & 0 & 0 \\ 0 & a_n(z) & \cdots & a_0(z) & \ddots & & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & & \ddots & \ddots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_n(z) & \cdots & a_0(z) \\ b_m(z) & \cdots & b_0(z) & 0 & \cdots & 0 & 0 \\ 0 & b_m(z) & \cdots & b_0(z) & \ddots & & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & & \ddots & \ddots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & b_m(z) & \cdots & b_0(z) \end{pmatrix}.$$

The coefficient vector $(a_0(z), \ldots, a_n(z)) \in \mathcal{M}(\Omega)^{n+1}$ *is shifted* $(m-1)$ *times to the right and the coefficient vector* $(b_0(z), \ldots, b_m(z)) \in \mathcal{M}(\Omega)^{m+1}$ $(n-1)$ *times to the right. Sometimes the degrees of the polynomials will be indicated by writing* $R_{n,m}(p, q)$ *instead of* $R(p, q)$.

- The determinant of $R_{n,m}(p, q)$ is zero if and only if p and q have a common factor of positive degree or $a_n = b_m = 0$.

Definition 3.13 (Discriminant matrix of $p(s, a(z))$). *The discriminant matrix of $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ is defined by $D_n(p) = R(p, p') = R_{n,n-1}(p, p') \in \mathcal{M}(\Omega)^{(2n-1)\times(2n-1)}$, where $p'(s, a(z)) = \sum_{k=1}^{n} k a_k(z) s^{k-1}$.*

- The determinant of the discriminant matrix of a polynomial $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ is zero if and only if $p(s, a(z))$ and $p'(s, a(z))$ have a non-trivial common factor or if $a_n = 0$.

Definition 3.14 (Critical point of a polynomial). *Let $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ be a polynomial with meromorphic coefficients and with a first coefficient not identically zero, i.e. $a_n(z) \not\equiv 0$.*

The point $z_0 \in \Omega$ is called a critical point of $p(s, a(z))$ if and only if one of the following conditions is satisfied.

- *The first coefficient $a_n(z)$ vanishes at z_0.*
- *The point z_0 is a pole of one of the coefficients $a_i(z) \quad i \in \{0, \ldots, n\}$.*
- *The polynomial $p(s, a(z_0)) \in \mathbb{C}[s]$ has a strictly smaller number of distinct roots than $p(s, a(z))$ for some other $z \in \Omega$.*

The set of critical points of the polynomial $p(s, a(z))$ is denoted by C_p.

Definition 3.15 (Critical point of a matrix). *If the entries of $A(z)$ depend analytically on one complex parameter z in a domain $\Omega \subseteq \mathbb{C}$, then $z_0 \in \Omega$ is a critical point of $A(\cdot) = (A(z))_{z\in\Omega}$ if and only if z_0 is a critical point of the characteristic polynomial $\xi_{A(z)}(s) = \det(sI - A(z))$.*

The set of critical points of $A(\cdot)$ is denoted by C_A.

Now, the first important theorem of this section can be stated. It concludes about the structure of the critical points and it provides a factorization of the polynomial $p(s, a(z))$.

Theorem 3.16 (Critical points are isolated)**.** *Let $p(s, a(z)) \in \mathcal{M}(\Omega)[s]$ be a polynomial with meromorphic coefficients and with a first coefficient $a_n(z) \not\equiv 0$, and let D be a simply connected subset of $\Omega \backslash C_p$.*

It follows that

- *C_p is locally finite in Ω , i.e. $K \cap C_p$ is finite for compact $K \subseteq \Omega$, that*
- *there exist $n = deg(p)$ not necessarily distinct analytic functions $s_1(z), \ldots, s_n(z) \in \mathcal{O}(D)$, such that*

$$p(s, a(z)) = a_n(z) \prod_{i=1}^{n} (s - s_i(z)) \qquad \forall s \in D.$$

  and that
- *the multiplicity of each root $s_i(z)$ is constant on D.*

Proof. In a first step, the theorem will be proved for irreducible monic polynomials. Subsequently, the theorem will be proved for the general case where $p(s, a(z)) = a_n(z) \prod_{i=1}^{l} p_i^{m_i}(s, b^i(z))$ is factorized in irreducible monic factors $p_i(s, b^i(z)) \in \mathcal{M}(\Omega)[s]$.

Step 1: Proof for irreducible monic polynomials.

Let $q(s, b(z)) \in \mathcal{M}(\Omega)[s]$ be an irreducible monic polynomial of degree m and let $b(z) = (b_0(z), \ldots, b_{m-1}(z), 1)$ be its coefficient vector.

Step 1.a: $C_q = Z_q \cup P_q$ is locally finite.

The set P_q of all poles of the coefficients $b_i(z)$ of $q(s, b(z))$ is locally finite because the poles of a meromorphic function which isn't identically ∞ do not have a cluster point in a domain.

The set Z_q of multiple roots of $q(s, b(z))$ is locally finite as well. Since $q(s, b(z))$ is irreducible, $q(s, b(z))$ cannot have a non-constant common factor with $\frac{d}{ds}(q(s, b(z)))$ in $\mathcal{M}(\Omega)[s]$. Therefore, the discriminant of $q(s, b(z))$,

$\psi_q(z) = D_m(q)(z) \in \mathcal{M}(\Omega)$, must be a non-identically-zero meromorphic function on Ω.

$\psi_q(z)$ is a product of meromorphic functions and therefore meromorphic on $\Omega \backslash P_q$. For each point $z_0 \in \Omega \backslash P_q$, $\psi_q(z_0)$ is the discriminant of the complex polynomial $q(s, b(z_0)) \in \mathbb{C}[s]$. The aforementioned discriminant is zero in z_0 if and only if $q(s, b(z_0))$ has multiple roots.

Step 1.b: Analytic representation and constant multiplicity.

By theorem 3.6,for every z_0 in a simply connected domain $D \subseteq \Omega \backslash C_q$, there exists a neighborhood U of z_0 in D and m distinct analytic functions $s_1(\cdot), \ldots, s_m(\cdot) : U \to \mathbb{C}$ such that $q(s, b(z)) = \prod_{i=1}^{m} (s - s_i(z))$ and $s_i(z) \neq s_j(z)$ for $i \neq j, z \in U$. It follows from the monodromy theorem that the representation of $q(s, b(z))$ as a product of distinct analytic linear factors can be extended across the whole simply connected domain D.

Step 2: General case

The above-mentioned result will now be applied to the monic factors of $p(s, a(z)) = a_n(z) \prod_{i=1}^{l} p_i^{m_i}(s, b^i(z))$.

Step 2.a: C_p is locally finite.

Step 2.a.1: $P_p \cup \Omega_0$ is locally finite and $\bigcup_{i=1}^{l} P_{p_i} \subseteq P_p$

The set of poles P_p of the coefficients $a_i(z), i \in \{0, \ldots, n\}$, of the polynomial $p(s, a(z))$ and the set Ω_0 of zeros of the leading coefficient $a_n(z)$ are evidently locally finite. Furthermore, it is easy to see that every pole z_0 of a coefficient of $p_i(s, b^i(z))$ is a pole of one of the coefficients of $p(s, a(z)) = a_n(z) \prod_{i=1}^{l} p_i^{m_i}(s, b^i(z))$. Hence, the set of poles P_{p_i} of the polynomial $p(s, b^i(z))$ is for every $i \in \{1, \ldots, l\}$ contained within P_p.

Step 2.a.2: $Z = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{l} Z_{i,j}$ is locally finite.

$Z = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{l} Z_{i,j}$ denotes the set of common zeros of the polynomials p_i and p_j

for all $i \neq j$. Since the polynomials $p_i(s, b^i(z))$ and $p_j(s, b^j(z))$ with respective degrees n_i and n_j are coprime for every $i, j \in \{0, \ldots, l\}$ and $i \neq j$, the determinant of the resultant matrix $\psi_{i,j}(z) = \det\left[R_{n_i,n_j}(p_i, p_j)\right] \in \mathcal{M}(\Omega)$ cannot be identically zero on Ω. $\psi_{i,j}(z)$ is analytic on $\Omega \backslash P_p$. For each point $z_0 \in \Omega \backslash P_p$, $\psi_{i,j}(z_0)$ is the determinant of the resultant matrix of the complex polynomials $p_i(s, b^i(z_0))$ and $p_j(s, b^j(z_0))$. The function $\psi_{i,j}(z)$ is zero for a $z_0 \in \Omega \backslash P_p$ if and only if $p_i(s, b^i(z_0))$ and $p_j(s, b^j(z_0))$ have common roots. The set $Z_{i,j}$ of zeros of $\psi_{i,j}(z)$ in $\Omega \backslash P_p$ is locally finite in Ω, and therefore $Z = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{l} Z_{i,j}$ is also locally finite.

Step 2.a.3: $C_p = \Omega_0 \cup P_p \cup Z \cup \bigcup_{i=1}^{l} Z_{p_i}$.

If $z_0 \in \Omega \backslash (\Omega_0 \cup P_p \cup Z \cup \bigcup_{i=1}^{l} Z_{p_i})$, where Z_{p_i} denotes the set of multiple roots of $p_i(s, b^i(z))$, the complex polynomial $p(s, a(z_0))$ has $\nu = \sum_{i=1}^{l} n_i$ distinct roots. This is the maximal number of distinct roots of $p(s, a(z))$ for $z \in \Omega \backslash (\Omega_0 \cup P_p)$. On the other hand if $z_0 \in Z_{p_i}$, i.e. $p_i(s) = p_i(s, b^i(z_0))$ has strictly less than n_i distinct roots, or if $z_0 \in Z_{i,j}$ for some $i, j \in \{1, \ldots, l\}, i \neq j$, then $p(s, a(z_0))$ has strictly less than ν distinct roots. Thus

$$C_p = \Omega_0 \cup P_p \cup Z \cup \bigcup_{i=1}^{l} Z_{p_i}$$

such that C_p (as a finite union of locally finite sets) is locally finite in Ω.

Step 2.b: Analytic representation and constant multiplicity.

Let D be a simply connected subset of $\Omega \backslash C_p$. Since $C_{p_i} = P_{p_i} \cup Z_{p_i} \subseteq C_p$ for $i \in \{1, \ldots, l\}$, all the irreducible factors p_i can be decomposed on D into linear factors with analytical roots. Hence, there exist n (not necessarily distinct) analytic functions $s_i(\cdot)$ from D to $\mathbb{C}$ such that the equation $p(s, a(z)) = a_n(z) \prod_{i=1}^{n} (s - s_i(z))$ holds for every s in D. It follows from the definition of critical points that two roots $s_i(z)$ and $s_j(z)$ of $p(s, a(z))$ which coincide at some point $z \in D \subseteq \Omega \backslash C_p$ coincide across the whole domain D. ☐

Corollary 3.17. *If $A : \Omega \to \mathbb{C}^{n \times n}$ is an analytic matrix function on a domain $\Omega \subseteq \mathbb{C}$ it follows that*

- *the set of critical points C_A of $A(\cdot)$ C_A is locally finite,*
- *and, if $D \subseteq \Omega \backslash C_A$ is simply connected, that there exist n analytic functions $\lambda_i(\cdot) \in \mathcal{O}(D), i \in \{1, \ldots, n\}$, such that*

$$\Lambda(A(z)) = (\lambda_1(z), \ldots, \lambda_n(z)), \qquad \forall z \in D \text{ and}$$

that the multiplicity of each eigenvalue $\lambda_i(z)$ is constant on D.

Proof. This corollary follows directly from theorem 3.16 applied on the characteristic polynomial $\xi_A(s)$ of $A(\cdot)$. □

3.2 Behavior of the roots near critical points

The next theorem describes the behavior of the roots of a polynomial in the neighborhood of a critical point. The punctured disk $U_r^o(z_0)$ and the cut disk $U_r^-(x)$ are defined as

$$U_r^o(z_0) := U_r(z_0) \backslash \{z_0\}$$
$$U_r^-(x) := \{z \in U_r^o(x) | 0 < \arg(z - x) < 2\pi\}.$$

Cycles, which are obtained by analytical continuation of the roots, will play a crucial role in the remainder of this chapter.

Theorem 3.18 (Cycles)**.** *Let $p(s, a(z)) \in \mathcal{O}(\Omega)[s]$ be a monic polynomial, $U_r(z_0) \subseteq \Omega$ a disk around the critical point z_0 with a sufficiently small radius r such that $U_r(z_0) \cap C_p = \{z_0\}$ and let l be the number of different roots of $p(s, a(z_0))$.*

The κ_j different roots of $p(s, a(z))$, $z \in U_r^o(z_0)$, which converge for $z \to z_0$ towards s_j can be partitioned into h_j sets $K_{j,k}$, $j \in \{1, \ldots, h_j\}$, of roots with the same multiplicity and with respective cardinality q_{jk} on $U_r^o(z_0)$, i.e.

$$\{1, \ldots, \kappa_j\} = \{\{K_{j,1}\}, \{K_{j,2}\}, \ldots, \{K_{j,h_j}\}\}.$$

The set $\{s_{jki}(z) \mid k \in \{1, \ldots, h_j\}, i \in \{1, \ldots, q_{jk}\}, z \in U_\delta^o(z_0)\}$ is called the s_j-group of roots of $p(s, a(z))$.

The q_{jk}-valued function $(s_{jk1}(\cdot), \ldots, s_{jkq_{jk}}(\cdot))$ on $U_\delta^o(z_0)$ obtained by analytic continuation of a root $s_{jki}(z)$ of $p(s, a(z))$ along a circular path around z_0 in

$U_r^o(z_0)$ is called a cycle.

These cycles have the following properties.

- *If $\mathbf{q_{jk} = 1}$, the function $s_{jk1}(z)$ can be continued analytically across all of $U_r(z_0)$.*
- *If $\mathbf{q_{jk} \geq 2}$, $(s_{jk1}(z), \ldots, s_{jkq_{jk}}(z))$ defines a q_{jk}-valued function on $U_r^o(z_0)$ with a so called branch point z_0 of the order $q_{jk} - 1$.*
- *The branches of these functions are represented by a Puiseux series of the form*

$$s_{jkm}(z) = \sum_{t=0}^{\infty} \alpha_{jkt} \left[e^{i\frac{2\pi}{q_{jk}}m} (z - z_0)^{\frac{1}{q_{jk}}} \right]^t, \qquad z \in U_r^-(z_0),$$

  where $j \in \{1, \ldots, l\}$, $k \in \{1, \ldots, h_j\}$, $l \in \{1, \ldots, q_{jk}\}$ and $m \in \{1, \ldots, q_{jk}\}$.
- *All the roots of the s_j-group converge for $z \to z_0$ towards the root s_j of $p(s, a(z_0))$. s_j is called the cent re of the cycle, i.e. $\lim_{z \to z_0} s_{jkl}(z) = \alpha_{jk0} = s_j$ for $j \in \{1, \ldots, l\}$, $k \in \{1, \ldots, h_j\}$ and $l \in \{1, \ldots, q_{jk}\}$.*

Proof. Step 1: First factorization of the polynomial.
Let $s_1, \ldots s_l$ denote the distinct roots of $p(s, a(z_0))$. Let $m_1, \ldots m_l$ be the corresponding multiplicities of $s_1, \ldots, s_l$. It follows from theorem 3.3 that for an $\varepsilon > 0$ which renders the closed disks with cent re s_i mutually disjoint a $\delta(\varepsilon)$ in $(0, r)$ can be chosen such that the polynomial $p(s, a(z))$ for every $z \in U_{\delta(\varepsilon)}(z_0)$ has the same number of zeros m_i in the disk $U_\varepsilon(s_i)$ as $p(s, a(z_0))$.

Since $U_{\delta(\varepsilon)}(z_0) \subseteq U_r(z_0)$, it follows that no other critical point than z_0 is contained within $U_{\delta(\varepsilon)}(z_0)$ and that therefore the number of distinct roots κ_j of $p(s, a(z))$ in $U_\varepsilon(s_j)$ is constant for all $z \in U_{\delta(\varepsilon)}^o(z_0)$.

On the simply connected cut disk $U_r^-(z_0)$, there exist a total of $\sum_{j=1}^{l} \kappa_j$ analytic functions $s_i(z)$, representing the roots of $p(s, a(z))$. The corresponding multiplicities of these roots are constant on $U_r^-(z_0)$.

For $z \in U_{\delta(\varepsilon)}^o(z_0) \cap U_r^-(z_0) = U_{\delta(\varepsilon)}^-(z_0)$, $s_i(z)$ belongs to exactly one of the disks $U_\varepsilon(s_j)$.

Grouping the roots on $U_r^-(z_0)$ accordingly to the various $U_\varepsilon(s_j)$ to which belong the roots for $z \in U_{\delta(\varepsilon)}(z_0)$ leads to the factorization

$$p(s, a(z)) = \prod_{j=1}^{l} \prod_{k=1}^{\kappa_j} (s - s_{jk}(z))^{\mu_{jk}}, \qquad \forall z \in U_r^-(z_0).$$

This factorization satisfies the following points.

- $\lim_{z \to z_0} s_{jk}(z) = s_j, \qquad k \in \{1, \ldots, \kappa_j\},\ j \in \{1, \ldots, l\}$
- μ_{jk} denotes the constant multiplicity of the root $s_{jk}(z)$ of $p(s, a(z))$ on $U_r^-(z_0)$.
- $\sum_{k=1}^{\kappa_j} \mu_{jk} = m_j$.

STEP 2: ANALYTIC CONTINUATION OF THE ROOTS s_{jk}.

An arbitrarily chosen $s_{jk}(z)$, where $j \in \{1, \ldots, l\}$ and $k \in \{1, \ldots, \kappa_j\}$ can be continued analytically along every arc in $U_r^o(z_0)$

When the function element $s_{jk}(z)$ in a small neighborhood of $z_1 = z_0 + \rho e^{-i\alpha}$, for $\rho \in (0, r)$ and $\alpha \in (0, 2\pi)$ fixed, is continued analytically along the circular arc

$$\gamma(t) = z_0 + \rho e^{it}, \qquad t \in [-\alpha, 2\pi - \alpha],$$

a new function element $\widetilde{s_{jk}}(z)$ at z_1 is obtained.

STEP 2.A: THE FACTORIZATION IS PRESERVED UNDER ANALYTIC CONTINUATION.

Due to the identity theorem, the factorization of $p(s, a(z)) = \prod_{j=1}^{l} \prod_{k=1}^{\kappa_j} (s - s_{jk}(z))^{\mu_{jk}}$ is preserved by analytic continuation. This is because $p(s, a(z)) = \widetilde{p(s, a(z))} = \prod_{j=1}^{l} \prod_{k=1}^{\kappa_j} (s - \widetilde{s_{jk}}(z))^{\widetilde{\mu_{jk}}}$ for all $z \in U_r^-(z_0)$.

STEP 2.B: THE ROOTS REMAIN IN $U_\varepsilon(s_j)$.

When $s_{jk}(z)$ is continued at some point in $U_\delta^-(z_0)$ along a circular arc in $U_\delta^o(z_0)$, the resulting root element $\widetilde{s_{jk}}(z)$ remains in $U_\varepsilon(s_j)$.

- The resulting root element $\widetilde{s_{jk}}(z)$ coincides with one of the analytic roots $s_{jk'}(z), k' \in \{1, \ldots, \kappa_j\}$.

- Since multiplicities remain constant on the simply connected set $U_r^-(z_0)$ it follows that $\widetilde{\mu_{jk}} = \mu_{jk'} = \mu_{jk}$. Therefore, the multiplicity of the resulting root element $\widetilde{s_{jk}(z)}$ is the same one as that of the analytically continued root $s_{jk}(z)$.

Step 3: Second factorization.

When the application $\pi_{z_0} : s_{jk}(\cdot) \mapsto s_{jk'}(\cdot)$ denotes analytic continuation along a circular arc around z_0, not only the roots associated to one circle $U_\varepsilon(s_j)$ remain invariant but also the set of roots with identical multiplicity.

Therefore, the set $\{1, \ldots, \kappa_j\}$ can be partitioned into $\{\{K_{j,1}\}, \ldots, \{K_{j,h_j}\}\}$. h_j is the number of different multiplicities. The sets $\{K_{j,i}\}$, $i \in \{1, \ldots, h_j\}$, contain the indexes of roots in $U_\varepsilon(s_j)$ with identical multiplicity.

The factorization of $p(s, a(z))$, taking those invariance groups into account, is

$$p(s, a(z)) = \prod_{j=1}^{l} \prod_{k=1}^{h_j} \prod_{i=1}^{q_{jk}} (s - s_{jki}(z))^{\mu_{jk}}, \qquad z \in U_r^-(z_0),$$

where q_{jk} is the cardinality of $K_{j,k}$, i.e. $q_{jk} = |K_{j,k}|$.

For every $j \in \{1, \ldots, l\}$ and $k \in \{1, \ldots, h_j\}$ $s_{jkm}(z)$ is analytically continued to $s_{jk(m+1)}(z)$, $m \in \{1, \ldots, q_{jk} - 1\}$ and $s_{jkq_{jk}}(z)$ is analytically continued to $s_{jk1}(z)$ on $U_r^o(z_0)$.

Step 4: Representation of cycles by Puiseux series

The q_{jk}-valued function $f(z)$ which belongs to the partition $K_{j,k}$ and to the cycle $(s_{jk1}(z), \ldots, s_{jkq_{jk}}(z))$ associates with every point z in the punctured disk $U_r^o(z_0)$ a tuple of q_{jk} distinct values of the following form.

$$f(z) = \begin{cases} \big(s_{jk1}(z), \ldots, s_{jkq_{jk}}(z)\big) & \text{, if } z \in U_r^-(z_0) \\ \left(\lim\limits_{t \nearrow 2\pi} s_{jki}(z_0 + \rho e^{it}) | i \in \{1, \ldots, q_{jk}\}\right) & \text{, if } z = z_0 + \rho,\ 0 < \rho < r \end{cases}$$

By continuity, $s_j = \lim\limits_{z \to z_0} s_{jk}(z)$ holds for every $k \in \{1, \ldots, \kappa_j\}$. This limit is called cent re of the cycle.

- If $q_{jk} = 1$, the point $z = z_0$ is a removable singularity of $f(z)$ such that $f(z)$ defines an analytic function on the whole disk $U_r(z_0)$.

- If $q_{jk} \geq 2$, z_0 is called a branch point of order $(q_{jk}-1)$. The root functions $s_{jkm}(z), m \in \{1, \ldots, q_{jk}\}$ are called the branches of $f(z)$ on $U_r^-(z_0)$.

Step 4.a: Construction of the series.

Firstly, set $\zeta = (z - z_0)^{\frac{1}{q_{jk}}}$ and define two sets

$$\widetilde{U_r}^o(0) = \{\zeta \in \mathbb{C} \ : \ 0 < |\zeta| < r^{\frac{1}{q_{jk}}}\} \text{ and}$$
$$C_m = \{s \in \widetilde{U_r}^o(0) \ : \ \tfrac{m-1}{q_{jk}} 2\pi \leq \arg(s) < \tfrac{m}{q_{jk}} 2\pi\}, \quad k \in \{1, \ldots, q_{jk}\}.$$

It follows that the application

$$\varphi : \begin{cases} \widetilde{U_r}^o(0) & \rightarrow \quad U_r^o(z_0) \\ \zeta & \mapsto \quad z_0 + \zeta^{q_{jk}} \end{cases}$$

maps every sector C_m bijectively on $U_r^o(z_0)$.

The point ζ_1 in the sector C_1 is now considered. The point ζ_1 is in the pre-image

$$\varphi^{-1}(z_1) = \{\zeta_n = \rho^{\frac{1}{q}} e^{-\frac{i\alpha}{q} + \frac{n-1}{q} 2\pi} \ : \ n \in \{1, \ldots, q_{jk}\}\}$$

of a point $z_1 = z_0 + \rho e^{-i\alpha}$, which in turn is in the cut disk $U_r^-(z_0)$. The map

$$\zeta \xrightarrow{\varphi} \varphi(\zeta) \xrightarrow{s_{jk1}} s_{jk1}(\varphi(\zeta))$$

is analytic near ζ_1. It determines a power series $S_1(\zeta)$ at ζ_1. The continuation of $S_1(\zeta)$ along all paths in $\widetilde{U_r}^o(0)$ is noted $F(\zeta)$.

Step 4.b: $F(\zeta)$ is single-valued.

As ζ travels along a circular arc $\tilde{\gamma}$ in the ζ - plane once around 0, $\varphi(\zeta) = z_0 + \zeta^q$ travels q times around z_0 in $U_r^o(z_0)$. Therefore, $S_1(\zeta) = s_{jk1}(\varphi(\zeta))$ is analytically continued via $S_2(\zeta) = s_{jk2}(\varphi(\zeta))), \ldots, S_{q_{jk}}(\zeta) = s_{jkq_{jk}}(\varphi(\zeta))$, to $S_{q_{jk}+1}(\zeta) = S_1(\zeta)$. Hence, $S_1(\zeta) = s_{jk1}(\varphi(\zeta))$ remains unchanged by analytic continuation along $\tilde{\gamma}$ and $F(\zeta)$ coincides with $s_{jki}(z)$ on the open sector $\text{int}(C_i)$.

Step 4.c: Laurent series expansion.

The Laurent series expansion of $F(\zeta)$ is a priori

$$F(\zeta) = \sum_{k=-\infty}^{\infty} \alpha_k \zeta^k \text{ on } \widetilde{U_r}^o(0).$$

However, since $\lim_{\zeta \to 0} F(\zeta) = \alpha_0 = s_j$, it follows that all coefficients α_k are zero for $k < 0$.

Thus,

$$F(\zeta) = s_j + \sum_{k=1}^{\infty} \alpha_k \zeta^k, \qquad \zeta \in U_{r^{\frac{1}{q}}}(0)$$

is analytic and satisfies

$$F(\zeta) = s_{jkm}(\varphi(\zeta)), \qquad \zeta \in \operatorname{int}(C_m),\ m \in \{1, \ldots, q_{jk}\}.$$

The q_{jk} values of f at $z \in U_r^o(z_0)$ are given by $F(\zeta_\nu)$, $\nu \in \{1, \ldots, q\}$, where ζ_ν runs through all the q_{jk}-th roots of $(z - z_0)$, i.e.

$$f(z) = \left\{ F(\zeta_\nu) \mid \zeta_\nu = \rho^{\frac{1}{q_{jk}}} e^{i\frac{\nu}{q_{jk}}2\pi + i\frac{\theta}{q_{jk}}} \ \wedge\ \nu \in \{1, \ldots, q_{jk}\} \right\},$$

where $z = z_0 + \rho e^{i\theta} \in U_r^o(z_0)$ and $0 \le \theta < 2\pi$. □

The same result can also be established for the characteristic polynomial of an analytic matrix function $A(z)$.

Corollary 3.19. *Let $A : \Omega \to \mathbb{C}^{n \times n}$ be an analytic matrix function on a domain $\Omega \subseteq \mathbb{C}$. Let $U_r(z_0) \subseteq \Omega$ be a disk around the critical point $z_0 \in C_A$ with a sufficiently small radius r such that $U_r(z_0) \cap C_A = \{z_0\}$ and let l be the number of different eigenvalues of $A(z)$.*

The κ_j different eigenvalues of $A(z)$, $z \in U_r^o(z_0)$, which converge for $z \to z_0$ towards λ_j, can be partitioned into h_j sets of roots with the same multiplicity and with respective cardinality q_{jk} on $U_r^o(z_0)$, i.e.

$$\{1, \ldots, \kappa_j\} = \{\{K_{j,1}\}, \{K_{j,2}\}, \ldots, \{K_{j,h_j}\}\}.$$

The set $\{\lambda_{jki}(z) \mid k \in \{1, \ldots, h_j\}, i \in \{1, \ldots, q_{jk}\}, z \in U_\delta^o(z_0)\}$ is called the λ_j-group of eigenvalues of $A(z)$.

The q_{jk}-valued function $(\lambda_{jk1}(\cdot), \ldots, \lambda_{jkq_{jk}}(\cdot))$ on $U_\delta^o(z_0)$ obtained by analytic continuation of a root $\lambda_{jki}(z)$ of $p(s, a(z))$ along a circular path around z_0 in $U_r^o(z_0)$ is called a cycle.

These cycles have the following properties.

- *If $\mathbf{q_{jk} = 1}$, the function $\lambda_{jk1}(z)$ can be continued analytically across $U_r(z_0)$.*

- *If* $\mathbf{q_{jk}} \geq \mathbf{2}$*,* $(\lambda_{jk1}(z), \ldots, \lambda_{jkq_{jk}}(z))$ *defines a* q_{jk}*-valued function on* $U_r^o(z_0)$ *with a so called branch point* z_0 *of the order* $q_{jk} - 1$*.*

- *The branches of these functions are represented by a Puiseux series of the form*

$$\lambda_{jkm}(z) = \sum_{t=0}^{\infty} \alpha_{jkt} \left[e^{i\frac{2\pi}{q_{jk}}m} (z - z_0)^{\frac{1}{q_{jk}}} \right]^t, \qquad z \in U_r^-(z_0),$$

where $j \in \{1, \ldots, l\}$*,* $k \in \{1, \ldots, h_j\}$ *and* $m \in \{1, \ldots, q_{jk}\}$*.*

- *All the different eigenvalues in the* λ_j*-group converge for* $z \to z_0$ *towards the eigenvalue* λ_j *of* $A(z_0)$*.* λ_j *is called the cent re of the cycle, i.e.* $\lim_{z\to z_0} \lambda_{jkl}(z) = \alpha_{jk0} = \lambda_j$ *for* $j \in \{1, \ldots, l\}$*,* $k \in \{1, \ldots, h_j\}$ *and* $m \in \{1, \ldots, q_{jk}\}$*.*

Proof. This corollary follows directly from theorem 3.18 applied on the characteristic polynomial of $A(\cdot)$. □

3.3 Continuous representation of the root functions

If $z_0 \in C_p$ is a branch point for the roots of p, that is that the permutation π_{z_0} defined by the analytic continuation of the roots along a circular path around z_0 is not the identity, it is not possible to find continuous (single-valued) functions $s_i(z), i \in \{1, \ldots, n\}$ representing the complete set of roots $p(s, a(z))$ on a disk around the branch point z_0. However, if the parameter z is restricted to a real interval, it is possible to find a continuous parametrization of the roots. In this case, it is sufficient to assume the continuous dependence of the coefficient vector on $z = r$ in order to prove the continuous dependence of the roots. The following theorem is taken from [11].

Theorem 3.20 (Real case)**.** *If* $I \subseteq \mathbb{R}$ *is an interval in* $\mathbb{R}$ *and* $p(s, a(r)) \in \mathcal{C}(I, \mathbb{C})[s]$ *is a monic polynomial of degree* n*, then there exist* n *continuous functions* $s_i : I \to \mathbb{C}$ *such that*

$$\Lambda(a(r)) = \lfloor s_1(r), \ldots, s_n(r) \rfloor, \qquad \forall r \in I.$$

Proof. A subinterval I_0 of I has the property (A) if there exist n continuous functions on I_0 representing the roots of the polynomial $p(s, a(r))$.

Step 1: Property (A) is local.

Let $\lfloor s_1^{(1)}(r), \ldots, s_n^{(1)}(r) \rfloor$ and $\lfloor s_1^{(2)}(r), \ldots, s_n^{(2)}(r) \rfloor$ be representations of $\Lambda(a(r))$ in I_1 and I_2 respectively. It is assumed that

- $I_1 \not\subseteq I_2 \wedge I_2 \not\subseteq I_1$ and
- I_1 lies to the left of I_2.

For $r_0 \in I_1 \cap I_2$ and after suitable renumbering $s_i^{(1)}(r_0) = s_i^{(2)}(r_0), i \in \{1, \ldots, n\}$ it follows that $\lfloor s_1^{(1)}(r_0), \ldots, s_n^{(1)}(r_0) \rfloor$ and $\lfloor s_1^{(2)}(r_0), \ldots, s_n^{(2)}(r_0) \rfloor$ represent the same $\Lambda(a(r_0))$.

The functions $s_i^{(0)}(r), i \in \{1, \ldots, n\}$, are defined on I_0 by

$$s_i^{(0)}(r) = \begin{cases} s_i^{(1)}(r), & r \leq r_0 \\ s_i^{(2)}(r), & r \geq r_0. \end{cases}$$

These functions are continuous and represent $\Lambda(a(r))$ on I_0.

If every point r in a subinterval I' of I has a neighborhood U with property (A), then the whole subinterval I' has property (A). The property (A) is local.

Step 2: Proof by an induction argument.

For the case $n = 1$, the claimed representation is evident. To conclude in the case $(n-1) \mapsto n$, it will be shown that, for every point in a open subset of I, there exists a neighborhood which has property (A). This relation is equivalent to the fact that the open subset has property (A).

The set

$$\Theta = \{ r \in I \mid n \text{ elements of } \Lambda(a(r)) \text{ are identical } \}$$

is a pre-image of a singleton and therefore closed in I. Its complement

$$\Xi = I \backslash \Theta$$

is open in I.

Step 2.a: Every r in Ξ has a neighborhood U with property (A)

Since n elements of $\Lambda(a(r_0))$, $r_0 \in \Xi$, are not all identical, they can be divided into two separate groups with n_1 and n_2 elements, where $n_1 + n_2 = n$. In other words, $\Lambda(a(r_0))$ is composed of an n_1-tuple and an n_2-tuple. The elements in the n_1-tuple are different to the elements in the n_2-tuple; there is no element in the n_1-tuple of the same value as an element in the n_2-tuple.

STEP 2.B: $\Lambda(a(r))$ IS CONTINUOUS ON $U_\varepsilon(r_0)$.

Since $\Lambda(a(r))$ consists of an n_1-tuple an n_2-tuple for sufficiently small $|r - r_0|$, it follows on from the induction hypothesis that these tuples can be represented by families of continuous functions in a neighborhood $U_\varepsilon(r_0)$ of r_0.

The combination of the two tuples represents $\Lambda(a(r))$ in $U_\varepsilon(r_0)$. Hence, $U_\varepsilon(r_0)$ has property (A).

STEP 2.C: REPRESENTATION OF $\Lambda(a(r))$.

Since Ξ is open in I, I consists, at most, of a countable number of subintervals $I_1, I_2, \ldots$. Since every $r \in \Xi$ has a neighborhood $U_\varepsilon(r)$ with property (A), each component interval I_p has property (A).

The n continuous functions $s_i(r)$, $i \in \{1, \ldots, n\}$, which represent $\Lambda(a(r))$ on I are defined by

$$s_i(r) = \begin{cases} s_i^{(p)}(r), & r \in I_p \wedge p \in \{1, 2, \ldots\} \\ s(r) & r \in \Theta. \end{cases}$$

□

Again, the theorem can be established for matrices and the corresponding eigenvalues too.

Corollary 3.21 (Real case)**.** *If $I \subseteq \mathbb{R}$ is an interval in $\mathbb{R}$ and $A(\tau) \in \mathbb{C}^{n \times n}$ depends continuously on τ in I, then there exist n continuous functions $\lambda_i : I \to \mathbb{C}$ such that*

$$\Lambda(A(\tau)) = \lfloor \lambda_1(\tau), \ldots, \lambda_n(\tau) \rfloor \qquad \forall \tau \in I.$$

Proof. Apply theorem 3.20 to the characteristic polynomial $\xi_{A(\tau)}(s) = det(sI - A(\tau))$. □

If $A(\tau)$ is differentiable in τ on an interval $I \subseteq \mathbb{R}$, and if $A(\tau)$ is diagonalizable for all $\tau \in I$, then, by a theorem of Kato, the $\lambda_i(\cdot)$ can be chosen to be differentiable.

4 Smoothness of Eigenprojections and Eigenvectors

In this section, the behavior of eigenprojections and eigenvectors for the case of not necessarily distinct eigenvalues will be investigated. It will be shown that the eigenprojections have necessarily poles on branch points. Furthermore, the normality condition, which prohibits the existence of any two root functions with the same multiplicity (i.e. branch points can not occur), will be introduced. In particular, this condition is satisfied by an self-adjoint operator.

4.1 Some facts about resolvents and eigenprojections of matrices

Definition 4.1 (Resolvent of A at s)**.** *Let A be a matrix in $\mathbb{C}^{n\times n}$ and $\rho(A) = \mathbb{C}\backslash\sigma(A)$ the resolvent set of A. Then*

$$R(s, A) = (sI - A)^{-1}, \quad s \in \rho(A)$$

is called the resolvent of A at $s \in \rho(A)$.

Some facts about resolvents will now be derived.

Lemma 4.2.

- *$R(s, A)$ commutes with A.*
- *The resolvent equation*

  $$R(s, A) - R(s_0, A) = (s_0 - s)R(s, A)R(s_0, A)$$

  holds for all s, s_0 in $\rho(A)$.
- *$R(s, A)$ and $R(s_0, A)$ commute.*
- *$R(s, A)$ admits the absolutely convergent power series expansion*

  $$R(s, A) = \sum_{k=0}^{\infty}(s - s_0)^k (R(s_0, A))^{k+1}, \qquad |s - s_0| < \|R(s_0, A)\|^{-1}$$

for every $s_0 \in \rho(A)$. Therefore, $R(s,A)$ is analytic on $\rho(A)$ and its derivatives at $s_0 \in \rho(A)$ are

$$R^{(k)}(s_0, A) = k!(R(s_0, A))^{k+1}, \qquad k \in \mathbb{N}^*.$$

Proof. STEP 1: $R(s,A)$ COMMUTES WITH A. Since the two equations

$$(sI - A)^{-1}(sI - A) = s(sI - A)^{-1} - (sI - A)^{-1}A = I$$
$$(sI - A)(sI - A)^{-1} = s(sI - A)^{-1} - A(sI - A)^{-1} = I$$

hold, it follows that

$$A(sI - A)^{-1} = (sI - A)^{-1}A.$$

STEP 2: RESOLVENT EQUATION.

If $s \neq s_0$, it follows that

$$\begin{aligned} R(s,A) - R(s_0,A) &= R(s,A)\underbrace{R(s_0,A)(s_0I - A)}_{=I} - \underbrace{(sI - A)R(s,A)}_{=I}R(s_0,A) \\ &= \underbrace{-R(s,A)R(s_0,A)A + AR(s,A)R(s_0,A)}_{\substack{=0, \\ \text{(because } A \text{ and } R(s,A) \text{ commute)}}} + (s_0 - s)R(s,A)R(s_0,A). \end{aligned}$$

STEP 3: $R(s,A)$ AND $R(s_0,A)$ COMMUTE.

Together with the resolvent equation, it follows from

$$\begin{aligned} R(s,A) - R(s_0,A) &= \underbrace{(s_0I - A)R(s_0,A)}_{=I}R(s,A) - R(s_0,A)\underbrace{R(s,A)(sI - A)}_{=I} \\ &= \underbrace{-AR(s_0,A)R(s,A) + R(s_0,A)R(s,A)A}_{\substack{=0, \\ \text{(because } A \text{ and } R(s,A) \text{ commute)}}} + (s_0 - s)R(s_0,A)R(s,A), \end{aligned}$$

that $R(s,A)$ and $R(s_0,A)$ commute.

STEP 4: POWER SERIES REPRESENTATION.

The resolvent equation is equivalent to

$$\begin{aligned} R(s_0,A) &= R(s,A) + (s_0 - s)R(s,A)R(s_0,A) \\ &= R(s,A)[I - (s - s_0)R(s_0,A)], \qquad s, s_0 \in \rho(A). \end{aligned}$$

Therefore, the following absolutely convergent series expansion of $R(s,A)$ at $s_0 \in \rho(A)$

$$R(s,A) = \sum_{k=0}^{\infty}(s - s_0)^k(R(s_0,A))^{k+1}, \qquad |s - s_0| < \|R(s_0,A)\|^{-1}$$

is obtained.

Thus, $R(s,A)$ is analytic on $\rho(A)$ and its derivatives at $s_0 \in \rho(A)$ are

$$R^{(k)}(s_0,A) = k!(R(s_0,A))^{k+1}, \qquad k \in \mathbb{N}^*.$$

□

In the following, $\lambda_1, \dots \lambda_l$ denote the distinct eigenvalues of $A \in \mathbb{C}^{n\times n}$ and $m_1, \dots, m_l$ their corresponding multiplicities.

$\mathbb{C}^n$ can be decomposed into A-invariant generalized eigenspaces.

$$\mathbb{C}^n = \ker(\lambda_1 I - A)^{m_1} \oplus \cdots \oplus \ker(\lambda_l I - A)^{m_l}.$$

Definition 4.3 (Eigenprojection P_j of A)**.** *The map*

$$P_j : \begin{cases} \mathbb{C}^n & \rightarrow \quad \ker(\lambda_j I - A)^{m_j} \\ x_1 \oplus \cdots \oplus x_l & \mapsto \quad x_j \end{cases}, \qquad j \in \{1, \dots, l\}$$

is called the eigenprojection of A for the eigenvalue λ_j.

Definition 4.4 (Eigennilpotent N_j of A)**.** *The operator*

$$N_j = (A - \lambda_j I)P_j : \mathbb{C}^n \rightarrow \mathbb{C}^n$$

is called the eigennilpotent of A for the eigenvalue λ_j.

Properties

(i). $\sum_{j=1}^{l} P_j = I, \qquad P_j P_k = \delta_{jk} P_j, \qquad \bowtie \in \{1, \dots, l\}$

(ii). $P_j A = A P_j = \lambda_j P_j + (A - \lambda_j I)P_j = \lambda_j P_j + N_j$

(iii). $A \sum_{j=1}^{l} P_j = \sum_{j=1}^{l} (\lambda_j P_j + N_j)$

(iv). A is diagonalizable if and only if $N_j = 0$, $j \in \{1, \dots, l\}$.

(v). If A is diagonalizable, i.e. $A = \sum_{j=1}^{l} \lambda_j P_j$, it follows that the resolvent can be written as $R(s,A) = \sum_{j=1}^{l} (s - \lambda_j)^{-1} P_j$ for every $s \in \rho(A)$

(vi). If A is normal, then A is diagonalizable. It also follows that the eigenprojections P_j are self-adjoint (i.e. $P_j = P_j^*$, $j \in \{1, \dots, l\}$) and that the spectral norm of P_j is 1 (i.e. the biggest eigenvalue is 1).

(vii). If A is real and symmetric, then the eigenvalues $\lambda_j, j \in \{1, \dots, l\}$, are real and the eigenprojections are real and symmetric.

4.2 Integral representation of the eigenprojections

Lemma 4.5 (Partial fraction expansion of R(s,A))**.** *Let $\lambda_1, \dots \lambda_l$ be the l distinct eigenvalues of $A \in \mathbb{C}^{n\times n}$. Let $m_1, \dots, m_l$ be the corresponding multiplicities. Let $P_1, \dots, P_l$ be the associated eigenprojections. And let $N_1, \dots, N_l$ be the associated eigennilpotents.*

It then follows that R(s,A) can be

$$R(s, A) = \sum_{j=1}^{l} \left[\frac{P_j}{(s-\lambda_j)} + \sum_{k=1}^{m_j-1} \frac{N_j^k}{(s-\lambda_j)^{k+1}} \right]. \tag{8}$$

If Γ_j is a positively oriented circle in $\rho(A)$ enclosing exclusively λ_j, it follows that the eigenprojection can be written as the integral

$$P_j = \frac{1}{2\pi i} \int_{\Gamma_j} R(s, A) ds. \tag{9}$$

Proof. Step 1: Without loss of generality, A has Jordan canonical form.

It can be easily seen that if $\tilde{A} = T^{-1}AT$, where $T \in GL_n(\mathbb{C})$ is a non singular matrix, the equations

$$R(s, \tilde{A}) = T^{-1}R(s, A)T, \qquad \tilde{P}_j = T^{-1}P_jT, \text{ and } \tilde{N}_j = T^{-1}N_jT,$$

hold as well for every j in $\{1, \dots, l\}$.

Therefore, A can be written as

$$A = \bigoplus_{j=1}^{l} \bigoplus_{k=1}^{r_j} J(\lambda_j, n_{jk}),$$

where $J(\lambda_j, n_{jk})$ is defined as

$$J(\lambda_j, n_{jk}) = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda_j & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_j \end{pmatrix}, \qquad \in \mathbb{C}^{n_{jk}\times n_{jk}}, k \in \{1, \dots, r_j\}.$$

In the above expression, r_j is the geometric multiplicity of the eigenvalue λ_j, n_{jk} the dimension of the k-th Jordan block of the eigenvalue λ_j and $m_j = \sum_{k=1}^{r_j} n_{jk}$ the algebraic multiplicity of the eigenvalue λ_j.

The associated eigenprojections P_j and eigennilpotents are,

$$\begin{aligned} P_j &= 0_{m_1} \oplus \cdots \oplus 0_{m_{j-1}} \oplus I_{m_j} \oplus 0_{m_{j+1}} \oplus \cdots \oplus 0_{m_l} \\ N_j &= 0_{m_1} \oplus \cdots \oplus 0_{m_{j-1}} \oplus \sum_{k=1}^{r_j} J(0, n_{jk}) \oplus 0_{m_{j+1}} \oplus \cdots \oplus 0_{m_l}, \end{aligned}$$

where I_n denotes an identity matrix of dimension n and 0_n a matrix of zeros of dimension $(n \times n)$.

Step 2: Formula for R(s,A).

Since A has Jordan normal form, the resolvent can be written as

$$R(s, A) = \bigoplus_{j=1}^{l} \bigoplus_{k=1}^{r_j} R(s, J(\lambda_j, n_{jk})).$$

Furthermore, for any $\lambda \in \mathbb{C}$, $m \in \mathbb{N}$,

$$\begin{aligned} R(s, J(\lambda, m)) &= \begin{pmatrix} (s-\lambda) & -1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ 0 & & & (s-\lambda) \end{pmatrix}^{-1} = \\ &= \frac{1}{\det(R(s, J(\lambda, m)))} \left(\widehat{R(s, J(\lambda, m))} \right) = \\ &= \begin{pmatrix} (s-\lambda)^{-1} & (s-\lambda)^{-2} & \cdots & (s-\lambda)^{-m} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-\lambda)^{-2} \\ 0 & \cdots & 0 & (s-\lambda)^{-1} \end{pmatrix} \\ &= (s-\lambda)^{-1} I_m + \sum_{k=1}^{m-1} (s-\lambda)^{-k-1} N^k. \end{aligned}$$

Step 3: Integral representation of P_j.

By Cauchy's formula,

$$\frac{1}{2\pi i} \int_\gamma \frac{f(s)}{(s-\lambda)^{k+1}} ds = \left. \frac{f^{(k)}(s)}{k!} \right|_{s=\lambda}.$$

In the above expression, γ is a positively oriented circle around λ and $f(z)$ an analytic function. It follows that

$$\frac{1}{2\pi i}\int_{\Gamma_j} R(s, J(\lambda_i, n_{ik}))ds = \begin{cases} 0_{n_{ik}} & \text{if } i \neq j \\ 1_{n_{ik}} & \text{if } i = j \end{cases}$$

Since

$$P_j = 0_{m_1} \oplus \cdots \oplus 0_{m_{j-1}} \oplus 1_{m_j} \oplus 0_{m_{j+1}} \oplus \cdots \oplus 0_{m_l},$$

this proves the assertion. □

Corollary 4.6. *Let $\lambda_1, \dots \lambda_l$ denote the l distinct eigenvalues of $A \in \mathbb{C}^{n\times n}$ and let Γ be a positively oriented simple closed curve in $\rho(A)$, enclosing the eigenvalues of A. It follows that the equation*

$$\frac{1}{2\pi i}\int_{\Gamma} R(s, A)ds = \sum_{j=1}^{l} P_j$$

holds.

Proof. Due to the residue theorem, it follows that

$$\frac{1}{2\pi i}\int_{\Gamma} R(s, A)ds = \sum_{j=1}^{l} \mathrm{Res}(R(s, A), \lambda_j) = \sum_{j=1}^{l} \frac{1}{2\pi i}\int_{\Gamma_j} R(s, A)ds = \sum_{j=1}^{l} P_j,$$

where $\Gamma_j, j \in \{1, \dots, k\}$ are small circles around λ_j. Thus, the assertion follows from corollary 4.6. □

4.3 Analyticity at non-critical points

In this section, the analyticity of the eigenprojections and the eigenbasis is investigated. Firstly, the total projection will be defined in order to understand the behavior of the projections when small changes occur. Secondly, the obtained results will be applied to the case where the matrices depend analytically on a complex parameter. Lastly, the analytical dependence of the eigenvectors will be proved.

4.3.1 Analyticity with respect to matrices

Definition 4.7 (λ_j - group of eigenvalues of $A = A_0 + \triangle$)**.** *Let $\lambda_1, \dots \lambda_l$ be the distinct eigenvalues of $A_0 \in \mathbb{C}^{n \times n}$ and let $m_1, \dots, m_l$ be the corresponding multiplicities. Furthermore, Γ_j, $j \in \{1, \dots, l\}$, denote the non-overlapping circles around each λ_j such that for all $\triangle \in \mathbb{C}^{n \times n}$ with a sufficiently small norm, Γ_j encloses exactly m_j eigenvalues of $A = A_0 + \triangle$, taking account of multiplicities.*

The set (or the unordered m_j-tuple) of the eigenvalues of $A = A_0 + \triangle$ is called λ_j - group of A.

Definition 4.8 (Total projection of the λ_j - group)**.** *Under the same framework as the preceding definition, the total projection for the λ_j-group of eigenvalues of $A = A_0 + \triangle$,*

$$P_j^{total}(A) = \frac{1}{2\pi i} \int_{\Gamma_j} R(s, A) ds, \quad \|A - A_0\| < \min_{s \in \Gamma_j} \|R(s, A_0)\|^{-1}.$$

- It follows from corollary 4.6 that the total projection coincides with the sum of eigenprojections of all the eigenvalues of A lying inside Γ_j.
- $P_j(A_0)$ coincides with the eigenprojection for the eigenvalue λ_j of A_0.

Proposition 4.9. *Let $\lambda_1, \dots \lambda_l$ be the distinct eigenvalues of $A_0 \in \mathbb{C}^{n \times n}$. Let $m_1, \dots, m_l$ be the corresponding multiplicities. Let $\|A - A_0\|$ be smaller than $\|R(s, A_0)\|^{-1}$. Then, it follows that the total projections $P_j(A)$ depend analytically on the entries of A.*

Proof. Step 1: Power series representation of $R(s, A)$.
$\triangle \in \mathbb{C}^{n \times n}$ is defined by $A - A_0 = \triangle$, where $\|\triangle\| < \min_{s \in \Gamma_j} \|R(s, A_0)\|^{-1}$ for all $j \in \{1, \dots, l\}$. Since

$$\begin{aligned} R(s, A) &= (sI - (A_0 + \triangle))^{-1} = \\ &= [(sI - A_0) - \triangle \underbrace{(sI - A_0)^{-1}(sI - A_0)}_{=I}]^{-1} = \\ &= [(I - \triangle \underbrace{(sI - A_0)^{-1}}_{=R(s, A_0)})(sI - A_0)]^{-1} = \\ &= (sI - A_0)^{-1}[I - \triangle R(s, A_0)]^{-1}, \end{aligned}$$

it follows that the power series

$$R(s,A) = R(s,A_0)\sum_{k=0}^{\infty}[(A-A_0)R(s,A_0)]^k$$

converges uniformly for $s \in \Gamma_j$ and for every $j \in \{1,\ldots,l\}$.

Step 2: Power series representation of $P_j(A)$ in the entries of $(A - A_0)$.

With the power series representation of $R(s,A)$, the formula

$$P_j(A) = \frac{1}{2\pi i}\sum_{k=0}^{\infty}\int_{\Gamma_j} R(s,A_0)[(A-A_0)R(s,A_0)]^k ds$$

is obtained easily. □

4.3.2 Analyticity with respect to a parameter

If $\|A - A_0\|$ is sufficiently small then A has at least as many distinct eigenvalues as A_0. On the other hand, every neighborhood of $A_0 \in \mathbb{C}^{n\times n}$ contains matrices with any number of different eigenvalues between $l = |\sigma(A_0)|$ and n. In the case of matrices depending analytically on a complex parameter, the situation greatly simplifies.

The following theorem, taken from [11], permits to define the eigenprojections on a simply connected domain using Kato's construction of globally defined transformation matrices $U(z, z_0)$.

Theorem 4.10 (Global definition of the transformation functions)**.** *Let $P_1(z),\ldots,P_n(z)$ be analytic projection-valued functions on a simply connected subset $D \subseteq \mathbb{C}$ and define $Q(z) = \frac{1}{2}\sum_{k=1}^{l}[P_k'(z)P_k(z) - P_k(z)P_k'(z)]$. The projections satisfy the following conditions.*

(i). $\sum_{j=1}^{l} P_j(z) = I_n$

(ii). $P_j(z)P_k(z) = \delta_{jk}P_k(z) \quad \forall j,k \in \{1,\ldots,n\}$

(iii). For every z_0 in D, $U(\cdot, z_0)$ is the unique solution of the operator differential equation $X'(z) = Q(z)X(z), X(z_0) = I$.

It then follows that the unique solution has the following properties.

(i). $U(z_0, z_0) = I$

(ii). $U(z_1, z)U(z, z_0) = U(z_1, z_0)$

(iii). $U(z, z_0)^{-1} = U(z_0, z)$

(iv). $P_j(z)$ *and* $P_j(z_0)$ *are similar, i.e.* $P_j(z)U(z, z_0) = U(z, z_0)P_j(z_0)$.

(v). *If* D *is symmetric with respect to the real axis, i.e. when* z *is an element of* D *then* $\overline{z}$ *is contained within* D *as well, and if the matrices* $P(z)$ *and* $P(z_0)$ *are self-adjoint, i.e.* $P(z)^* = P(\overline{z})$, $P(z_0)^* = P(\overline{z_0})$, *then* $U(\bar{z}, \bar{z_0})^* = U(z, z_0)^{-1}$.

In particular, if $z, z_0 \in D$ *are real and* $P(z)$ *and* $P(z_0)$ *are hermitian (resp. real and symmetric), then* $U(z, z_0)$ *is unitary (resp. orthogonal)*

Proof. By differentiating $P^2(z) = P(z)$ and $P_j(z)P_k(z) = \delta_{jk}P_k(z)$ it follows that

- $P_j'(z)P_j(z) + P_j(z)P_j'(z) = P_j'(z), \quad j \in \{1, \ldots, n\}$, and
- $P_j(z)P_k'(z) = P_j'(z)P_k(z), k \neq j$.

Step 1: $P_j'(z) = Q(z)P_j(z) - P_j(z)Q(z)$.

$$\begin{aligned}
\mathbf{Q(z)P_j(z)} &= \tfrac{1}{2}\sum_{k=1}^{l}\left[P_k'(z)\underbrace{P_k(z)P_j(z)}_{=\delta_{jk}P_j(z)} - \underbrace{P_k(z)P_k'(z)}_{P_k'(z)-P_k'(z)P_k(z)} P_j(z)\right] = \\
&= \tfrac{1}{2}P_j'(z)P_j(z) + \tfrac{1}{2}\sum_{k=1}^{l}\left[P_k'(z)\underbrace{P_k(z)P_j(z)}_{\delta_{jk}P_j(z)} - P_k'(z)P_j(z)\right] = \\
&= P_j'(z)P_j(z) - \tfrac{1}{2}\sum_{k=1}^{l}\underbrace{P_k'(z)P_j(z)}_{=\begin{cases} -P_k(z)P_j'(z) & k \neq j \\ -P_j(z)P_j'(z) + P_j'(z) & else \end{cases}} = \\
&= P_j'(z)P_j(z) + \tfrac{1}{2}\left[\left(\underbrace{\left(\sum_{k=1}^{l}P_k(z)\right)}_{=I}P_j'(z)\right) - P_j'(z)\right] = \\
&= \mathbf{P_j'(z)P_j(z)}
\end{aligned}$$

In the same manner one can conclude that $P_j(z)Q(z) = -P_j(z)P_j'(z)$. From these two equations it follows that

$$P_j'(z) = (P_j(z)^2)' = P_j'(z)P_j(z) + P_j(z)P_j'(z) = Q(z)P_j(z) - P_j(z)Q(z)$$

STEP 2: PROPERTIES OF THE TRANSFORMATION FUNCTION.

The first two claims about the evolution operator $U(\cdot, z_0)$ are known from the theory of operator differential equation. The third one is a consequence of the second by setting $z_1 = z_0$.

For the last equation, the following differential equations are considered.

$$\begin{aligned}(P_j(z)U(z,z_0))' &= P_j'(z)U(z,z_0) + P_j(z)\underbrace{U'(z,z_0)}_{=Q(z)U(z,z_0)}\\ &= (P_j'(z) + P_j(z)Q(z))U(z,z_0)\\ &= Q(z)P_j(z)U(z,z_0)\\ (U(z,z_0)P_j(z_0))' &= U'(z,z_0)P_j(z_0)\\ &= Q(z)U(z,z_0)P(z_0)\end{aligned}$$

Since the solution of the equation $X'(z) = Q(z)X(z)$ is unique, the equation $P_j(z)U(z,z_0) = U(z,z_0)P_j(z_0)$ must hold for every $j \in \{1, \ldots, n\}$. The initial value of this solution is $X(z_0) = P_j(z_0)$.

STEP 3: IF $P(z)$ AND $P(z_0)$ ARE SELF-ADJOINT, IT FOLLOWS THAT $U(z,z_0)^{-1} = U(\overline{z}, \overline{z_0})^*$.

Since the equations

$$\begin{aligned}\mathbf{Q}(\overline{\mathbf{z}})^* &= \frac{1}{2}\sum_{k=1}^{l}\left[P_k(\overline{z})^*P_k'(\overline{z})^* - P_k'(\overline{z})^*P_k(\overline{z})^*\right]\\ &= \frac{1}{2}\sum_{k=1}^{l}\left[P_k(z)P_k'(z) - P_k'(z)P_k(z)\right]\\ &= -\mathbf{Q}(\mathbf{z})\end{aligned}$$

and

$$\frac{d}{dz}U(\overline{z},\overline{z_0})^* = -U(\overline{z},\overline{z_0})^*Q(z)$$

hold, it follows that

$$\begin{aligned}\tfrac{d}{dz}(U(\overline{z},\overline{z_0})^*U(z,z_0)) &= U'(\overline{z},\overline{z_0})^*U(z,z_0) + U(\overline{z},\overline{z_0})^*U'(z,z_0) =\\ &= -U(\overline{z},\overline{z_0})^*Q(z)U(z,z_0) + U(\overline{z},\overline{z_0})^*Q(z)U(z,z_0) = 0.\end{aligned}$$

Therefore, the product $U(\overline{z},\overline{z_0})^*U(z,z_0)$ must be constant.

The initial conditions, for both differential equations at $z = z_0$ are $U(\overline{z_0}, \overline{z_0}) = I$ and $U(z_0, z_0) = I$ respectively. In view of this fact, it follows that

$$U(\overline{z}, \overline{z_0})^* U(z, z_0) = I \qquad \forall z \in D.$$

Given that the dimension of the underlying space is finite, this is equivalent to

$$U(\overline{z}, \overline{z_0})^* = U(z, z_0)^{-1}.$$

□

Corollary 4.11 (Analyticity of total projections)**.** *Let $A : \Omega \to \mathbb{C}^{n \times n}$ be analytic on the domain $\Omega \subseteq \mathbb{C}$, let $\lambda_1, \dots, \lambda_l$ be the l distinct eigenvalues of $A_0 = A(z_0)$ and let $m_1, \dots, m_l$ be the corresponding multiplicities. If $|z - z_0|$ is sufficiently small, then the total projection $P_j(z) := P_j(A(z))$ depends analytically on z and has constant rank, rank$(P_j(z)) = m_j, j \in \{1, \dots, l\}$.*

Proof. Since the total projection has a power series representation in the entries of $A(z)$ it depends analytically on z on a small disk. The rank is constant because $P_j(z)$ and $P_j(z_0)$ are similar due to theorem 4.10. □

Moreover, the total projection coincides on non critical points with the eigenprojection. These are analytic on any simply connected subset of Ω which does not contain critical points.

Corollary 4.12 (Analyticity of eigenprojections)**.** *Let $A : \Omega \to \mathbb{C}^{n \times n}$ be analytic on the domain $\Omega \subseteq \mathbb{C}$ and $\lambda_j : D \to \mathbb{C}, j \in \{1, \dots, l\}$, be the l distinct eigenvalue functions on the simply connected set $D \subseteq \Omega \backslash C_A$ such that $\sigma(A(z)) = \{\lambda_1(z), \dots, \lambda_l(z)\}$ for $z \in D$.*

Then, it follows that

- *the corresponding eigenprojections $P_j(z), j \in \{1, \dots, l\}$ are analytic and have constant rank, rank$(P_j(z)) = m_j, j \in \{1, \dots, l\}$, on D, that*
- *the eigennilpotents $N_j = (A(z) - \lambda_j(z)I)P_j(z), j \in \{1, \dots, l\}$ are analytic on D as well and that*

- *$A(z)$ admits the spectral representation*

$$A(z) = \sum_{j=1}^{l} (\lambda_j(z) P_j(z) + N_j(z)), \quad \forall z \in D.$$

Proof. STEP 1: $P_j(z)$ HAS CONSTANT RANK ON $D \subseteq \Omega \backslash C_A$.
The claim follows directly from theorem 4.10, because $P_j(z)$ and $P_j(z_0)$ are similar for $z, z_0 \in D$.

STEP 2: THE EIGENPROJECTION ARE ANALYTIC ON $D \subseteq \Omega \backslash C_A$.
Let $A_0 = A(z_0)$ be the matrix belonging to an arbitrary non-critical point $z_0 \in \Omega \backslash C_A$. Let $\Gamma_j, j \in \{1, \dots, l\}$, be non-overlapping circles around the $l = l(z_0)$ eigenvalues of A_0.

For every $z_0 \in D$ exists a disk with cent re z_0 such that

- there is exactly one eigenvalue $\lambda_j(z)$ of $A(z)$ inside the circle Γ_j and
- the total projection $P_j(z) = \dfrac{1}{2\pi i} \displaystyle\int_{\Gamma_j} R(s, A(z)) ds = \operatorname{Res}(R(s, A(z)), \lambda_j(z))$ is the eigenprojection for $\lambda_j(z)$.

Therefore, the analyticity of the eigenprojections follows from Proposition 4.9.

□

4.3.3 Analyticity and Construction of the generalized eigenbasis

Given an analytic eigenvalue $\lambda_j(\cdot)$ of algebraic multiplicity m_j on a subdomain D of $\Omega \backslash C_A$, it is of interest to find for each $z \in D$ m_j generalized eigenvectors $v^{j,k}(z), k \in \{1, \dots, m_j\}$. They form a basis for the generalized eigenspace $\ker(\lambda_j(z) I - A(z))^{m_j}$ and they depend analytically on $z \in D$.

Corollary 4.13. *Let $P : D \to \mathbb{C}^{n \times n}$ be an analytic projection-valued function on a simply connected set $D \subseteq \mathbb{C}$, let z and z_0 be points in D and let $(v^1, \dots, v^m)$ be a basis for $\operatorname{Im}(P(z_0)) = P(z_0)\mathbb{C}^n$. Furthermore, let $U(z, z_0)$ be the transformation function from theorem 4.10.*

Under these assumptions, the vectors $v^i(z) = U(z, z_0) v^i, i \in \{1, \dots, m\}$, form a basis of $\operatorname{Im}(P(z)) \subseteq \mathbb{K}^n$ which depends analytically on $z \in D$.

Proof. Since

$$\mathbf{P(z)v^i(z)} = P(z)U(z,z_0)v^i = U(z,z_0)P(z)v^i = U(z,z_0)v^i = \mathbf{v^i(z)}$$

the vectors $v^i(z), i \in \{1,\ldots,m\}$ belong to $\mathrm{Im}(P(z))$. Since $\dim(\mathrm{Im}(P(z))) = \dim(\mathrm{Im}(P(z_0)))$ and since $U(z,z_0)$ is invertible, the vectors $v^i(z), i \in \{1,\ldots,m\}$, are linearly independent and form for every $z \in D$ a basis of $\mathrm{Im}(P(z))$. □

Corollary 4.14. *Let $A : \Omega \to \mathbb{C}^{n\times n}$ be analytic on the domain $\Omega \subseteq \mathbb{C}$. Furthermore, λ_j is an eigenvalue of algebraic multiplicity m_j of $A(z_0)$, $z_0 \in \Omega\backslash C_A$. It follows that there exists*

- *an analytic function $\lambda_j(\cdot)$ on $U_\delta(z_0)$ representing an eigenvalue of algebraic multiplicity m_j of $A(z).\lambda_j(\cdot)$ coincides in z_0 with λ_j, i.e. $\lambda_j(z_0) = \lambda_j$.*
- *Furthermore, there are m_j analytic vector functions $(v^{j,1}(z),\ldots,v^{j,m_j}(z))$ from $U_\delta(z_0)$ to $\mathbb{C}^n$ constituting a basis of the generalized eigenspace $\ker(\lambda_j(z))I - A(z))^{m_j}$ of $A(z)$ for all z in $U_\delta(z_0)$.*

Proof. Step 1: $\lambda_j(z)$ is analytic eigenvalue function for $A(z)$.
Since z_0 is a non-critical point, the λ_j - group of eigenvalues of $A(z)$ near $z = z_0$ consists of only one eigenvalue $\lambda_j(z)$ of multiplicity m_j. This eigenvalue of $A(z)$ depends analytically on z in a small disk $U_\delta(z_0)$.

Step 2: $(v^{j,1}(z),\ldots,v^{j,m_j}(z))$ is a basis of the generalized eigenspace.
The associated total projection $P_j(z)$ coincides with the eigenprojection for $\lambda_j(z)$ and depends analytically on $z \in U_\delta(z_0)$.

By corollary 4.13, it follows that the vectors $v^{j,k}(z) = U(z,z_0)v^{j,k}, k \in \{1,\ldots,m_j\}$, form a basis of the generalized eigenspace $\mathrm{Im}(P(z)) = \ker(\lambda_j(z)I - A(z))^{m_j}$ and that they depend analytically on $z \in U_\delta(z_0)$. □

4.4 Analytic properties of eigenprojections near critical points

Following the chapter about perturbed polynomials, the behavior of eigenprojections near critical points will be investigated. Even though the branch points of the eigenvalues and the eigenprojections are the same, the behavior of the eigenprojections differ considerably from the behavior of the eigenvalues.

In the following, $z_0 \in \Omega$ denotes a possibly critical point of $A(\cdot)$. $\lambda_1, \dots, \lambda_l$ denote the distinct eigenvalues of $A(z_0)$ and $m_1, \dots, m_j$ denote the corresponding algebraic multiplicities. Each non-overlapping positively oriented circle $\Gamma_j, j \in \{1, \dots, l\}$, encloses exactly m_j eigenvalues of $A(z)$. κ_j denotes the maximal number of distinct eigenvalues, enclosed by each Γ_j for $z \in U_\delta^o(z_0)$. These eigenvalues are called the λ_j-group of eigenvalues of $A(z)$ near z_0. Furthermore, the punctured disk $U_\delta^o(z_0) \subseteq \Omega$ does not contain any critical point of $A(\cdot)$.

In analogy to the preceding chapter, $(\lambda_{jk1}(z), \dots, \lambda_{jkq_{jk}}(z))$, $k \in \{1, \dots, h_j\}$, denote for every j in $\{1, \dots, l\}$ the h_j different cycles of eigenvalues of constant multiplicity $m_{jki}(z) = m_{jk}(z) = m_{jk}$, obtained by analytic continuation in $U_\delta^o(z_0)$.

Altogether, $A(z)$ has a total of $\sum\limits_{j=1}^{l} \kappa_j = \sum\limits_{j=1}^{l} \sum\limits_{k=1}^{h_j} q_{jk}$ distinct eigenvalues for every z in $U_\delta^o(z_0)$. Taking their multiplicities into account, they add up to a total number of $\sum\limits_{j=1}^{l} \kappa_j \sum\limits_{k=1}^{h_j} q_{jk} m_{jk} = n$ eigenvalues.

Theorem 4.15. *Let $A : \Omega \to \mathbb{C}^{n \times n}$ be an analytic matrix function on a domain $\Omega \subseteq \mathbb{C}$ and let $A(z_0) = \sum\limits_{j=1}^{l} \lambda_j P_j + N_j$ be the spectral representation of $A(z_0)$ at some (possibly critical) point $z_0 \in \Omega$.*

Under these assumptions, the following statements hold.

- *The eigenprojections and eigennilpotents of $A(z)$ are branches of analytic functions of $z \in \Omega$ with at most algebraic singularities at $z_0 \in C_A$ and their branch points are the same as those of the corresponding eigenvalues.*
- *On the punctured disk $U_\delta^o(z_0)$, $A(z)$ can be represented as*

$$A(z) = \sum_{j=1}^{l} \sum_{k=1}^{h_j} \sum_{i=1}^{q_{jk}} (\lambda_{jki}(z) P_{jki}(z) - N_{jki}(z)).$$

- *If* $\mathbf{q_{jk}=1}$, *the functions* $\lambda_{jk1}(z), P_{jk1}(z)$ *and* $N_{jk1}(z)$ *can be continued analytically across the whole disk* $U_\delta(z_0)$.
- *If* $\mathbf{q_{jk}\geq 2}$, *the branches of the eigenprojection cycle* $(P_{jk1}(z),\dots,P_{jkq_{jk}}(z))$ *near* z_0 *can be represented by a Laurent-Puiseux series of the form*

$$P_{jkm}(z)=\sum_{l=-l_{jk}}^{\infty} B_{jkl}(e^{i\frac{2\pi}{q_{jk}}m}(z-z_0)^{\frac{1}{q_{jk}}})^l,\quad z\in U_\delta^-(z_0),$$

where $l_{jk}\in\mathbb{N}$, $B_{jkl}\in\mathbb{C}^{n\times n}$ *for* $l\geq -l_{jk}$.

Proof. Step 1: Analyticity of the total projections.
The total projection $P_j(z)$, associated with one of the λ_j - groups

$$\{\lambda_{jkm}(z)\ :\ k\in\{1,\dots,h_j\}, m\in\{1,\dots,q_{jk}\}\}, z\in U_\delta^o,$$

is the sum of all the eigenprojections corresponding to the κ_j eigenvalues of the λ_j-group. Due to corollary 4.11, it depends analytically on z in a small disk with cent re z_0.

Step 2: Analyticity of the eigenprojections and eigennilpotents.
By corollary 4.12, the eigenprojection

$$P_{jki}(z):\bigoplus_{j'=1}^{l}\bigoplus_{k'=1}^{h_{j'}}\bigoplus_{i'=1}^{q_{j'k'}}\ker(\lambda_{j'k'i'}(z)I-A(z))^{m_{j'k'}}\to\ker(\lambda_{jki}(z)I-A(z))^{m_{jk}}$$

of $A(z)$, corresponding to the eigenvalue branch $\lambda_{jki}(z)$, is defined and analytic on the (simply connected) cut disk $U_\delta^-(z_0)$. The same statement holds for the associated eigennilpotents

$$N_{jkm}(z)=(A(z)-\lambda_{jkm}(z)I)P_{jkm}(z).$$

In analogy to the eigenvalue branches, the eigenprojection branches $P_{jki}(z)$ can be continued analytically across $U_\delta^o(z_0)$. They are linked one with another if and only if the same is true for the corresponding eigenvalue branches $\lambda_{jki}(z)$ because of the following reasoning.

Step 2.a: The permutation to which the families of eigenvalues, eigenprojections and eigennilpotents are subjected, are identical.
Consider the families $(\lambda_h(z))_{h\in\{1,\dots,c\}}$, $(P_h(z))_{h\in\{1,\dots,c\}}$ and $(N_h(z))_{h\in\{1,\dots,c\}}$

consisting of $\sum_{j=1}^{l} \kappa_j = c$ elements. The resolvent $R(s,z)$ of $A(z)$ is defined, in accordance with formula (8), as

$$R(s,z) = \sum_{h=1}^{c} \left[\frac{P_h(z)}{(s-\lambda_j(z))} + \sum_{k=1}^{m_j-1} \frac{N_h(z)^k}{(s-\lambda_j(z))^{k+1}} \right],$$

where s is assumed to be somewhere distant from the spectrum $\sigma(A(z))$ of $A(z)$ such that s is contained within $\rho(A(z))$ for $z \in U_\delta^o(z_0)$. Let $\lambda_1(z), \ldots, \lambda_p(z)$ be a cycle of eigenvalues. The permutation π_{z_0} maps $\lambda_h(z)$ to $\lambda_{h+1}(z)$, $h \in \{1, \ldots, p-1\}$, and $\lambda_p(z)$ to $\lambda_1(z)$. However, since the resolvent $R(s,z)$, which is unique as a meromorphic operator-valued function in s, remains unchanged by analytic continuation, the permutation must map $P_h(z)$ to $P_{h+1}(z)$, $h \in \{1, \ldots, p-1\}$, and $P_p(z)$ to $P_1(z)$. The possibility that $P_h(z) = P_k(z)$ for some $h \neq k$ is excluded by the property $P_h(z)P_k(z) = \delta_{hk}P_h(z)$. Similar results hold for the eigennilpotents $N_h(z)$, except that some pair in the set $\{N_h(z) \mid h \in \{1, \ldots, c\}\}$ may coincide, since they could all be zero.

Therefore, the operator $\sum_{i=1}^{q_{jk}} \lambda_{jki}(z)P_{jki}(z)$ associated with the eigenvalue cycle $(\lambda_{jk1}(z), \ldots, \lambda_{jkq_{jk}}(z))$ is invariant with respect to analytic continuation along circular arcs in $U_\delta^o(z_0)$. It can be extended analytically across the punctured disk.

The application

$$P_j(z) = \sum_{k=1}^{h_j} \sum_{i=1}^{q_{jk}} P_{jki}(z), \quad z \in U_\delta^o(z_0)$$

is the total projection associated with the λ_j - group, and

$$A(z) = \sum_{j=1}^{l} \sum_{k=1}^{h_j} \sum_{i=1}^{q_{jk}} (\lambda_{jki}(z)P_{jki}(z) - N_{jki}(z)), \quad z \in U_\delta^o(z_0)$$

is the spectral representation of $A(z)$.

Step 3: Singularities are at most algebraic.

It is evident that

$$\|P_{jki}(z)\| = \left\| \frac{1}{2\pi i} \int_{\Gamma_{jki}(z)} R(s, A(z))ds \right\| \leq r(\Gamma_{jki}(z)) \max_{s \in \Gamma_{jki}(z)} \|R(s, A(z))\|,$$

where $\Gamma_{jki}(z)$ is a circle around $\lambda_{jki}(z)$ and $r(\Gamma_{jki}(z))$ is the radius of this circle.

Furthermore,

$$\|R(s,A(z))\| = \|(sI-A(z))^{-1}\| \le \gamma \frac{\|sI-A(z)\|^{n-1}}{|\det(sI-A(z))|} \le$$
$$\le \gamma \frac{(\|A(z)\|+|s|)^{n-1}}{\left|\prod_{j=1}^{l}\prod_{k=1}^{h_j}\prod_{i=1}^{q_{jk}}(s-\lambda_{jki}(z))\right|},$$

where $\gamma > 0$ is a constant depending only on the chosen norm. The first inequality will be proved in step 3.a.

Since $\lambda_{jki}(z) \xrightarrow{z\to z_0} \lambda_j(z_0), k \in \{1,\dots,h_j\}, i \in \{1,\dots,q_{jk}\}$, the circles $\Gamma_{jki}(z)$ must be chosen smaller for smaller $|z|$ in order to ensure that they remain non-overlapping.

By choosing $r(\Gamma_{jki}(z)) = |z|^\alpha$ with an appropriate $\alpha > 0$, it can be ensured that

$$\left|\prod_{j=1}^{l}\prod_{k=1}^{h_j}\prod_{i=1}^{q_{jk}}(s-\lambda_{jki}(z))\right| \ge \gamma'|z|^{\alpha n}, \quad s \in \Gamma_{jki}(z), \gamma' > 0$$

Hence,

$$\begin{aligned}\|P_{jki}(z)\| &\le \gamma r(\Gamma_{jki}) \max_{s\in\Gamma_{jki}} \frac{(\|A(z)\|+|s|)^{n-1}}{\left|\prod_{j=1}^{l}\prod_{k=1}^{h_j}\prod_{i=1}^{q_{jk}}(s-\lambda_{jki}(z))\right|} \le \\ &\le \gamma|z|^\alpha \max_{s\in\Gamma_{jki}} \frac{(\|A(z)\|+|s|)^{n-1}}{\gamma'|z|^{\alpha n}} \le \\ &\le \beta\frac{1}{|z|^{(n-1)\alpha}},\end{aligned}$$

where $\beta > 0$ is a constant.

Step 3.a: The inequality $\|A^{-1}\| \le \frac{\gamma}{|\det(A)|}\|A\|^{n-1}$ holds.

Since all norms in a finite dimensional vector space are equivalent, the following inequalities are satisfied.

$$\begin{aligned}\gamma_1'\|\cdot\| \quad &\le \quad \|\cdot\|_{max} \quad \le \quad \gamma_2'\|\cdot\| \\ \gamma_1''|||\cdot||| &\le |||\cdot|||_{max} \le \gamma_2''|||\cdot|||,\end{aligned}$$

where $\|\cdot\|$ denotes an arbitrary norm of $\mathbb{K}^n$, $\|x\|_{max} = \|x_1b_1+\dots+x_nb_n\|_{max} = \max_{j\in\{1,\dots,n\}}|x_j|$ ($b_j, \in \{1,\dots,n\}$ is an orthonormal basis of $\mathbb{K}^n$) denotes the maximum norm, $|||\cdot|||$ denotes an arbitrary induced operator norm and $|||A|||_{max} = \max_{(j,k)\in\{1,\dots,n\}^2}|a_{jk}|$ denotes the maximum matrix norm, which isn't an operator norm.

Therefore ,

$$\begin{aligned}
\|u\| &= \|A^{-1}v\| \\
&\leq \frac{1}{|\det(A)|} \left|\left|\left| \left[\left((-1)^{k+l} S_{kl}(A) \right)_{k,l=1}^{n} \right]^T \right|\right|\right| \|v\| \\
&\leq \frac{1}{|\det(A)|} \frac{1}{\gamma_1''} \max_{(j,k)\in\{1,\ldots,n\}^2} \underbrace{|S_{kl}(A)|}_{=\sum_{\sigma\in\Pi_{n-1}} sign(\sigma) a_{i_1\sigma(i_1)}\cdots a_{i_{(n-1)}\sigma(i_{(n-1)})}} \|v\| \\
&\leq \frac{1}{|\det(A)|} \frac{1}{\gamma_1''} \max_{(j,k)\in\{1,\ldots,n\}^2} |a_{jk}|^{n-1}(n-1)!\|v\| \\
&\leq \frac{(n-1)!}{|\det(A)|} \frac{\gamma_2''}{\gamma_1''} |||A|||^{n-1}\|v\|,
\end{aligned}$$

where $S_{kl}(A)$ is the determinant of the $(n-1) \times (n-1)$-dimensional matrix that results from deleting row k and column j of A. Hence, $|||A^{-1}||| \leq \frac{\gamma}{|\det(A)|}|||A|||^{n-1}$, where $\gamma > 0$ is a constant.

□

Theorem 4.16 (Butler)**.** *If z_0 is a branch point of order $(q_{jk} - 1) \geq 1$ of an eigenvalue cycle $(\lambda_{jk1}(z), \ldots, \lambda_{jkq_{jk}}(z))$ of $A(z)$, it follows that*

- *the Laurent-Puiseux expansion of the associated eigenprojection $P_{jki}(z)$ in powers of $(z - z_0)^{\frac{1}{q}}$ contains negative powers. Therefore, $\|P_{jki}(z)\| \xrightarrow{z\to z_0} \infty$.*

Proof. $P_{jki}(z)$ belongs to the cycle $(P_{jk1}(z), \ldots, P_{jkq_{jk}}(z))$ of eigenprojections. Analytic continuation along a small circle around z_0 changes $P_{jki}(z)$ to $P_{jk(i+1)}(z), i \in \{1, \ldots, q_{jk} - 1\}$ and $P_{jkq_{jk}}$ to $P_{jk1}(z)$.

If the Laurent Puiseux expansion were not containing negative powers of $(z - z_0)^{\frac{1}{q}}$, it would follow that the limits

$$\lim_{z\to z_0} P_{jk(i+1)}(z) = \lim_{z\to z_0} P_{jki}(z) = B_{jk0}$$

exist. But since

- $\lim\limits_{z\to z_0} P_{jki}(z)P_{jk(i+1)}(z) = 0 = B_{jk0}B_{jk0}$ and
- $\lim\limits_{z\to z_0} P_{jki}(z)P_{jki}(z) = B_{jk0}B_{jk0} = \lim_{z\to z_0} P_{jki}(z) = B_{jk0}$

it follows that $B_{jk0}B_{jk0} = B_{jk0} = 0$.

This is a contradiction to the fact that $P_{jki}(z)$ is a non-zero projection and that $\|P_{jki}(z)\| \geq 1$ must hold. □

Therefore, analyticity of eigenprojections at branch points cannot be expected. A property of a matrix family, which ensures that branch points of $A(z)$ does not exist for $z \in \Omega$, is the following one.

Definition 4.17 (Normality condition at a point z_0)**.** *A continuous matrix function $A : \Omega \to \mathbb{C}^{n\times n}$ satisfies the normality condition at a point $z_0 \in \Omega$ if and only if there exists a sequence $(z_i)_{i\in\mathbb{N}^*} \in \Omega\backslash\{z_0\}$ which converges towards z_0 and for which $A(z_i)$ is normal for every $i \in \mathbb{N}^*$.*

By continuity it follows that under the normality condition, $A(z_0)$ must be normal as well. The next theorem provides an important analyticity result for eigenvalues and eigenprojections.

Theorem 4.18 (Rellich)**.** *Let $A : \Omega \to \mathbb{C}^{n\times n}$ be an analytic matrix function which satisfies the normality condition at $z_0 \in \Omega$. Its spectral representation at z_0 is*

$$A(z_0) = \sum_{i=1}^{l} \lambda_j P_j.$$

The eigenvalues λ_j have algebraic multiplicities m_j.Under these conditions, it follows that

- *there are $\kappa_j \leq m_j$ analytic eigenvalue functions*

$$\lambda_{j,k} : U_\delta(z_0) \to \mathbb{C},$$

  which represent the λ_j - group of eigenvalues of $A(z)$ on $U_\delta(z_0)$, $\{\lambda_{j,1}(z), \dots, \lambda_{j,\kappa_j}(z)\}$, and which coincide on z_0 with λ_j, i.e. $\lambda_{j,k}(z_0) = \lambda_j$ for every $k \in \{1, \dots, \kappa_j\}$, and that

- *there exist $\kappa_j \leq m_j$ analytic matrix functions*

$$P_{j,k} : U_\delta(z_0) \to \mathbb{C}^{n\times n}, \qquad k \in \{1, \dots, \kappa_j\}.$$

  - *Furthermore, $P_{j,k}(z)$ is the eigenprojections of $A(z)$ for the eigenvalue $\lambda_{j,k}(z), k \in \{1, \dots, \kappa_j\}$,*

- *the sum of all eigenprojections $P_{j,k}(z)$ belonging to the λ_j-group coincides at $z = z_0$ with P_j, i.e.* $\sum_{k=1}^{\kappa_j} P_{j,k}(z_0) = P_j$ *and*
- *each $P_{j,k}(z)$, $j \in \{1, \ldots, l\} \wedge k \in \{1, \ldots, \kappa_j\}$, has constant rank on $U_\delta(z_0)$.*

- *The spectral representation of $A(z)$ for $z \in U_\delta^o(z_0)$ is given by*

$$A(z) = \sum_{j=1}^{l} \sum_{k=1}^{\kappa_j} \lambda_{j,k}(z) P_{j,k}(z).$$

Proof. In accordance to the notations introduced at the beginning of section 4.4, the neighborhood $U_\delta(z_0) \subseteq \Omega$ contains at most one critical point, i.e. $U_\delta^o(z_0) \cap C_A = \emptyset$. The non-overlapping circles $\Gamma_j, j \in \{1, \ldots, l\}$, with cent re λ_j in $\mathbb{C}$ enclose exactly m_j eigenvalues of $A(z)$ for $z \in U_\delta(z_0)$.

Furthermore, $(z_i)_{i \in \mathbb{N}^*}$ is a sequence which converges towards z_0 and for which $A(z_i)$ is normal, for every $i \in \mathbb{N}^*$.

Step 1: The $\lambda(z)$ are analytic on $U_\delta(z_0)$

If z_0 is a non-critical point, every λ_j - group of eigenvalues of $A(z)$ contains only one, ($\kappa_j = 1$), eigenvalue $\lambda_{j,1}(z)$. This eigenvalue is analytic on the whole disk $U_\delta(z_0)$.

If z_0 is a critical point, $A(z_i)$ is normal and therefore the spectral norm of all $P_{j,k}(z_i), i \in \mathbb{N}^*$ is 1. Since $\lim_{i \to \infty} \|P_{j,k}(z_i)\|_2 = 1$, it follows by Butler's theorem that z_0 isn't a branch point for any $\lambda_{j,k}(z)$, i.e. $q_{jk} = 1$ for every $k \in \{1, \ldots, h_j\}$. Therefore, $\lambda_{j,k}(z)$ can be continued analytically across $U_\delta(z_0)$ by corollary 3.19.

Step 2: The eigenprojections $P_{j,k}(z)$ and eigennilpotents $N_{j,k}(z)$ are analytic on $U_\delta(z_0)$.

By corollary 4.12, the eigenprojections $P_{jk}(z)$ are analytic and they have constant rank on $U_\delta^-(z_0)$. Since z_0 isn't a branch point for any $\lambda_{j,k}(z)$, the associated eigenprojections $P_{j,k}(z)$ and eigennilpotent $N_{jk}(z) = (A(z) - \lambda_{j,k}(z)I)P_{j,k}(z)$, $j \in \{1, \ldots, l\}$, $k \in \{1, \ldots, \kappa_j\}$ can be extended analytically across $U_\delta(z_0)$.

STEP 3: $A(z)$ IS DIAGONALIZABLE.

If $A(z_i)$ is normal for all $i \in \mathbb{N}^*$, it follows that these matrices are diagonalizable and hence that all $N_{j,k}(z_i) = 0$ for all $i \in \mathbb{N}^*$. Since z_0 is the limit point of the sequence $(z_i)_{i\in\mathbb{N}^*}$, it follows by the identity theorem that $N_{j,k}(z)$ is identically zero on $U_\delta(z_0)$.

STEP 4: THE SUM OF ALL EIGENPROJECTIONS $P_{j,k}(z)$ BELONGING TO THE λ_j-GROUP COINCIDES AT $z = z_0$ WITH P_j.

By continuity, see corollary 4.6, the total projections $P_j(z) = \sum_{k=1}^{\kappa_j} P_{j,k}(z)$ converge to P_j as z converges versus z_0, i.e. $\sum_{k=1}^{\kappa_j} P_{j,k}(z_0) = P_j$. □

The next corollary is a generalization of Rellich's local theorem on simply connected subsets.

Corollary 4.19 (Rellich global)**.** *Let $A : \Omega \to \mathbb{C}^{n\times n}$ be an analytic matrix function on Ω and $D \subseteq \Omega$ a simply connected subset. $A(\cdot)$ satisfies the normality condition at each critical point in D.*

Under these assumptions, there exist

- *a total of N distinct analytic functions $\lambda_i : D \to \mathbb{C}$ and*
- *N analytic projection-valued functions $P_i : D \to \mathbb{C}^{n\times n}$, $i \in \{1, \ldots, N\}$,*

where N is the number of distinct eigenvalues of $A(z)$ at non-critical points in Ω.

Furthermore, the spectral representation of $A(z)$ is given for every $z \in D\backslash C_A$ by

$$A(z) = \sum_{i=1}^{N} \lambda_i(z)P_i(z), \qquad \forall z \in D\backslash C_A.$$

In particular, $A(z)$ is diagonalizable for all $z \in D$.

Proof. By Rellich's theorem, for every $z_0 \in D$ the following statements hold for a small disk $U^o_\delta(z_0) \subseteq D$.

- There are N analytic functions $\lambda_i : U_\delta(z_0) \to \mathbb{C}$ such that

$$\sigma(A(z)) = \{\lambda_1(z), \dots, \lambda_N(z)\}, \quad \forall z \in U_\delta(z_0).$$

- There exist N analytic projection-valued functions $P_i : U_\delta(z_0) \to \mathbb{C}^{n\times n}$ such that
  - for every $z \in U_\delta^o(z_0)$, $P_i(z)$ is the eigenprojection of $A(z)$ corresponding to the eigenvalue $\lambda_i(z), i \in \{1, \dots, N\}$, and
  - the spectral representation of $A(z)$ is

$$A(z) = \sum_{i=1}^{N} \lambda_i(z) P_i(z)$$

    for every $z \in U_\delta^o(z_0)$.

Analytic continuation on the whole simply connected subset. Each of the functions $\lambda_i(\cdot)$ and $P_i(\cdot)$ can be continued analytically along any arc in D in a way that $A(z) = \sum_{i=1}^{N} \lambda_i(z) P_i(z)$ is the spectral representation of $A(z)$ for every non-critical point on the arc. Since D is simply connected this defines, by the monodromy theorem, N complex functions $\lambda_i : D \longrightarrow \mathbb{C}$ and N analytic projection-valued functions $P_i : D \to \mathbb{C}^{n\times n}$ which satisfy the corollary. □

An analytic eigenbasis of $A(z)$ can now be constructed with Kato's global transformation function.

Corollary 4.20. *Let $A : \Omega \to \mathbb{C}^{n\times n}$ be an analytic matrix function on Ω and $D \subseteq \Omega$ a simply connected subset. $A(\cdot)$ satisfies the normality condition at each critical point.*

Then, there exists an analytic matrix function

$$V : D \to \mathbb{C}^{n\times n}$$

which satisfies the two following conditions.

- *$V(z)$ is invertible, i.e. $V(z) \in GL_n(\mathbb{C})$*
- *$V(z)^{-1} A(z) V(z)$ is diagonal for every $z \in D$*

Proof. By corollary 4.19, there are N analytic functions $\lambda_i : D \to \mathbb{C}$, $i \in \{1, \ldots, N\}$, and N analytic projection valued functions $P_j : D \to \mathbb{C}^{n\times n}$, $i \in \{1, \ldots, N\}$, on D such that $A(z) = \sum_{i=1}^{N} \lambda_i(z)P_j(z)$ on $D \backslash C_A$.

By theorem 4.10, for every projection-valued function $P_i(z), z \in U_\delta(z_o)$ exists an invertible transformation function $U_i(z, z')$ which depends analytically on $(z, z') \in D \times D$.

By corollary 4.13, the vectors $U_i(z, z_0)v^{i,l}, l \in \{1, \ldots, m_i\}$ form a basis of $\mathrm{Im}(P_i(z))$, whenever $(v^{i,1}, \ldots, v^{i,m_i})$ is a basis of $\mathrm{Im}(P_i(z_0))$, $\forall j \in \{1, \ldots, l\}$, $\forall k \in \{1, \ldots, \kappa_j\}$.

By corollary 4.19, the vectors

$$U_i(z, z_0)v^{i,l}, \quad \forall j \in \{1, \ldots, l\}, k \in \{1, \ldots, \kappa_j\}, l \in \{1, \ldots, m_i\}$$

form an eigenbasis of $A(z)$ on D. Choosing these n vectors as columns of $V(z)$, it follows that $V(z)^{-1}A(z)V(z)$ is a diagonal matrix, whose entries are the eigenvalues $\lambda_i(z)$ of $A(z)$. □

4.5 Real Case

In this last section, the real case will be treated. Therefore, $A : I \to \mathbb{C}^{n\times n}$ is analytic on an open interval $I \subseteq \mathbb{R}$ and can be expanded into a power series $A(\tau) = \sum_{k=0}^{\infty} A_k(\tau - \tau_0)^k$, which is absolutely convergent on a small interval $U^{\mathbb{R}}_{r(\tau_0)}(\tau_0)$.

The complex analytic extension $A(z) = \sum_{k=0}^{\infty} A_k(z - \tau_0)^k, z \in U^{\mathbb{C}}_{r(\tau_0)}(\tau_0)$ does not contain any non-real critical point, i.e. the number of distinct eigenvalues of $A(\tau)$ at non-real $z \in U^{\mathbb{C}}_{r(\tau_0)}(\tau_0)$ will be equal to the number of distinct eigenvalues of $A(\tau)$ at non-critical real points $\tau \in I$.

The union D of all these disks is a simply connected domain in $\mathbb{C}$ with $I = D \cap \mathbb{R}$. The analytic extension of $A(\cdot)$ to D will be denoted by $A(\cdot)$ as well. By construction, D does not contain any critical point of $A(\cdot)$ off the real axis.

Corollary 4.21. *Let $A : I \to \mathbb{C}^{n\times n}$ be an analytic matrix function on an open interval $I \subseteq \mathbb{R}$. The maximum number of distinct eigenvalues of $A(\cdot)$ is N.*

If the interval I does not contain any critical points of $A(\cdot)$, then there exists

- *N analytic functions* $\lambda_j : I \to \mathbb{C},\ j \in \{1, \dots, N\}$, *which represent the distinct eigenvalues of* $A(\tau)$ *for every* $\tau \in I$ *and*

- *lN analytic projection-valued functions* $P_j : I \to \mathbb{C}^{n \times n},\ j \in \{1, \dots, N\}$ *representing the corresponding eigenprojections of* $A(\tau)$ *for every* $\tau \in I$.

If $A(\cdot)$ *satisfies the normality condition at all critical points in* I, *then there exist again* N *analytic functions* $\lambda_j : I \to \mathbb{C}$ *and* N *analytic projection-valued functions* $P_j : I \to \mathbb{C}^{n \times n},\ j \in \{1, \dots, N\}$, *representing the eigenvalues and the corresponding eigenprojections of* $A(\tau)$ *for every* $\tau \in I$. *Furthermore,*

- *the spectral representation of* $A(\tau)$ *is given by*

$$A(\tau) = \sum_{j=1}^{N} \lambda_j(\tau) P_j(\tau), \quad \tau \in I$$

  and

- $A(\tau)$ *is diagonalizable for every* $\tau \in I$.

If $A(\tau)$ *is in addition self-adjoint, i.e.* $A(\tau) = A(\tau)^*$ *for* $\tau \in I$, *then*

- *the previous statement holds with self-adjoint projections* $P_j(\tau) = P_j(\tau)^* \in \mathbb{K}^{n \times n}$ *and*

- *there exists an analytic orthonormal basis of* $\mathbb{C}^n$ *consisting of eigenvectors of* $A(\cdot)$ *on* I, *i.e.*

$$V : I \to \mathbb{C}^{n \times n},$$

  where $V(\tau)$ *is orthogonal and* $V(\tau)^{-1} A(\tau) V(\tau)$ *is diagonal for every* $\tau \in I$.

Proof. $A(\cdot)$ is the at the beginning of this section constructed analytic extension across a simply connected domain $D \subseteq \mathbb{C}$ with $D \cap \mathbb{R} = I$.

Step 1: 1^{st} and 2^{nd} statement

Since the interval I does not contain any critical point, it follows that D does not contain any critical points. Therefore, the first statement follows from corollary 4.12 and corollary 3.17 and the second one from corollary 4.19.

Step 2: Analytic orthogonal eigenbasis.

Since the eigenprojections of a Hermitian matrix are self-adjoint, it follows that $P_i(\tau) = P_i(\tau)^*$, $i \in \{1, \ldots, N\}$, for all non-critical points $\tau \in I$ and hence by continuity that the equation is true for all $\tau \in I$. Furthermore, the eigenvectors of self-adjoint matrices are mutually orthogonal.

By the last part of theorem 4.10, it follows that an orthogonal matrix $U_i(\tau, \tau')$, $i \in \{1, \ldots, N\}$, which depends analytically on $(\tau, \tau') \in I \times I$ exists.

By corollary 4.13, the vectors $v^{ik}(\tau) = U_i(\tau, \tau')v^{ik}$, $k \in \{1, \ldots, m_i\}$ form an orthonormal basis of $\mathrm{Im}(P_i(\tau))$, whenever $(v^{i1}, \ldots, v^{im_i})$ is an orthonormal basis of $Im(P_i(\tau_0))$,$i \in \{1, \ldots, N\}$. This is because $U_i(\tau, \tau')$ is unitary.

Therefore, the vectors $v^{ik}(\tau)$, $i \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, m_i\}$, form an orthonormal basis of eigenvectors of $A(\tau)$ for all $\tau \in I$. □

5 Bibliography

[1] H. Baumgärtel, Analytic Perturbation Theory for Matrices and Operators, Birkhäuser, 1985.

[2] D. R. Brillinger, Time Series: Data Analysis and Theory, Holden-Day, San Francisco, 1981.

[3] J. B. Conway, Functions of one complex variable I, Springer, New York, 1978.

[4] M. Deistler and C. Zinner, Modelling High-Dimensional Time Series by Generalized Linear Dynamic Factor Models: An Introductory Survey, Communications in Information and Systems Vol.7, No.2, pp. 153-166, 2007.

[5] B.A. Fuks, Theory of analytic functions of several complex variables, American Mathematical Soc., Providence, 1965.

[6] I.M. Gelfand, D. A. Raikov, G. E. Shilov, Kommutative normierte Algebren, VEB Deutscher Verlag der Wissenschaften, 1964.

[7] Hamann E., Deistler M., Scherrer W., Factor Models for Multivariate Time Series, Adaptive information systems and modelling in economics and management science pp. 243 - 251, Wien, 2005.

[8] E. J. Hannan and M. Deistler, The Statistical Theory of Linear Systems, Wiley, New York, 1988.

[9] H. Havlicek, Lineare Algebra für technische Mathematiker, Heldermann, Lemgo, 2006.

[10] D. Hinrichsen, A. J. Pritchard, Mathematical Systems Theory I - Modelling, State Space Analysis, Stability and Robustness, Springer, Berlin, 2005.

[11] T. Kato, Perturbation theory for linear operators, Springer, Berlin, 1995.

[12] Y. A. Rozanov, Stationary Random Processes, Holden-Day, San Francisco, 1967.

[13] W. Rudin, Functional analysis, McGraw-Hill, New York, 1998.

[14] W. Scherrer and M. Deistler, A structure theory for linear dynamic errors-in-variables models, SIAM, J. on Control Optimisation, vol. 36(6), pp. 2148–2175, 1998.