1

Polynomial Matrices

1.1 Basic Concepts of Algebra

1. Let a certain set A with elements $a, b, c, d, \ldots$ be given. Assume that over the set A an algebraic operation is defined which relates every pair of elements (a, b) to a third element $c \in A$ that is called the result of the operation.

If the named operation is designated by the symbol '$*$', then the result is symbolically written as

$$a * b = c\,.$$

In general, we have $a * b \neq b * a$. However, if for any two elements a, b in A the equality $a * b = b * a$ holds, then the operation '$*$' is called *commutative*. The operation '$*$' is named *associative*, if for any $a, b, c \in A$ the relation

$$(a * b) * c = a * (b * c)$$

is true.

The set A is called a *semigroup*, if an associative operation '$*$' is defined in it. A semigroup A is called a *group* , if it contains a neutral element e, such that for every $a \in A$

$$a * e = e * a = a$$

is correct, and furthermore, for any $a \in A$ there exists a uniquely determined element $a^{-1} \in A$, such that

$$a * a^{-1} = a^{-1} * a = e\,. \tag{1.1}$$

The element a^{-1} is called the *inverse* element of a. A group, where the operation '$*$' is commutative, is called a commutative group or *Abelian group*.

In many cases the operation '$*$' in an Abelian group is called *addition*, and it is designated by the symbol '$+$'. This notation is called *additive*. In additive notation the neutral element is called the zero element , and it is denoted by the symbol '0' (zero).

In other cases the operation '$*$' is called *multiplication*, and it is written in the same way as the ordinary multiplication of numbers. This notation is named *multiplicative*. The neutral element in the multiplicative notation is designated by the symbol '1' (one). For the inverse element in multiplicative notation is used a^{-1}, and in additive notation we write $-a$. In the last case the inverse element $-a$ is also named the *opposite* element to a.

2. The set A is called an (associative) ring, if the two operations 'addition' and 'multiplication' are defined on A. Hereby, the set A forms an Abelian group with respect to the 'addition', and a semigroup with respect to the 'multiplication'. From the membership to an Abelian group it follows

$$(a+b)+c = a+(b+c)$$

and

$$a+b = b+a\,.$$

Moreover, there exists a zero element 0, such that for an arbitrary $a \in A$

$$a+0 = 0+a = a\,.$$

The element 0 is always uniquely determined. Between the operations 'addition' and 'multiplication' of a ring the relations

$$(a+b)c = ac+bc\,, \qquad c(a+b) = ca+cb$$

(left and right distributivity) are valid.

In many cases rings are considered, which possess a number of further properties. If for any two a, b always $ab = ba$ is true, then the ring is called *commutative*. If a unit element exists with $1a = a1$ for all $a \in A$, then the ring is named as a *ring with unit element*. The element 1 in such a ring is always uniquely determined.

The non-zero elements a, b of a ring, satisfying $ab = 0$, are named *(left resp. right) zero divisor*. A ring is called *integrity region* , if it has no zero divisor.

3. A commutative associative ring with unit element, where every non-zero element a has an inverse a^{-1} that satisfies Equation (1.1), is called a *field*. In others words, a field is a ring, where all elements different from zero with respect to multiplication form a commutative group. It can be shown that an arbitrary field is an integrity region. The set of complex or the set of real numbers with the ordinary addition and multiplication as operations are important examples for fields. In the following, these fields will be designated by $\mathbb{C}$ and $\mathbb{R}$, respectively.

1.2 Polynomials

1. Let $\mathcal{N}$ be a certain commutative associative ring with unit element, especially it can be a field. Let us consider the infinite sequence $(a_0, a_1, \ldots, a_k; 0, \ldots)$, where $a_k \neq 0$, and all elements starting from a_{k+1} are equal to zero. Furthermore, we write

$$(a_0, a_1, \ldots, a_k; 0, \ldots) = (b_0, b_1, \ldots, b_k; 0, \ldots)\,,$$

if and only if $a_i = b_i$ $(i = 0, \ldots, k)$. Over the set of elements of the above form, the operations addition and multiplication are introduced in the following way. The sum is defined by the relation

$$(a_0, a_1, \ldots, a_k; 0, \ldots) + (b_0, b_1, \ldots, b_k; 0, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots, a_k + b_k; 0, \ldots)$$

and the product of the sequences is given by

$$\begin{aligned} &(a_0, a_1, \ldots, a_k; 0, \ldots)(b_0, b_1, \ldots, b_k; 0, \ldots) \\ &\quad = (a_0 b_0, a_0 b_1 + a_1 b_0, \ldots, a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0, \ldots, a_k b_k; 0, \ldots)\,. \end{aligned} \tag{1.2}$$

It is easily proven that the above explained operations addition and multiplication are commutative and associative. Moreover, these operations are distributive too. Any element $a \in \mathcal{N}$ is identified with the sequence $(a; 0, \ldots)$. Furthermore, let λ be the sequence

$$\lambda = (0, 1; 0, \ldots)\,.$$

Then using (1.2), we get

$$\lambda^2 = (0, 0, 1; 0, \ldots)\,, \quad \lambda^3 = (0, 0, 0, 1; 0, \ldots)\,, \quad \textit{etc.}$$

Herewith, we can write

$$\begin{aligned} &(a_0, a_1, \ldots, a_k; 0, \ldots) = \\ &\qquad = (a_0; 0, \ldots) + (0, a_1; 0, \ldots) + \ldots + (0, \ldots, 0, a_k; 0, \ldots) \\ &\qquad = a_0 + a_1(0, 1; 0, \ldots) + \ldots + a_k(0, \ldots, 0, 1; 0, \ldots) \\ &\qquad = a_0 + a_1\lambda + a_2\lambda^2 + \ldots + a_k\lambda^k\,. \end{aligned}$$

The expression on the right side of the last equation is called a *polynomial* in λ with coefficients in $\mathcal{N}$. It is easily shown that this definition of a polynomial is equivalent to other definitions in elementary algebra. For $a_k \neq 0$ the polynomial $a_k\lambda^k$ is called the term of the polynomial

$$f(\lambda) = a_0 + a_1\lambda + \ldots + a_k\lambda^k \tag{1.3}$$

with the highest power. The number k is called the *degree* of the polynomial (1.3), and it is designed by $\deg f(\lambda)$. If we have in (1.3) $a_0 = a_1 = \ldots = a_k = 0$, then the polynomial (1.3) is named the *zero polynomial* . A polynomial with $a_k = 1$ is called *monic*. If for two polynomials $f_1(\lambda)$, $f_2(\lambda)$ the relation $f_1(\lambda) = af_2(\lambda)$ with $a \in \mathcal{N}$ is valid, then these polynomials are called *equivalent*. In what follows, we will use the notation $f_1\lambda) \approx f_2(\lambda)$ for the fact that the polynomials $f_1(\lambda)$ and $f_2(\lambda)$ are equivalent.

Inside this book we only consider polynomials with coefficients from the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$. Following [206] we use the notation $\mathbb{F}$ for a field that is either $\mathbb{R}$ or $\mathbb{C}$. The set of polynomials over these fields are designated by $\mathbb{R}[\lambda]$, $\mathbb{C}[\lambda]$ or $\mathbb{F}[\lambda]$ respectively. The sets $\mathbb{R}[\lambda]$ and $\mathbb{C}[\lambda]$ are commutative rings without zero divisor. In what follows, the elements in $\mathbb{R}[\lambda]$ are called real polynomials.

2. Some general properties of polynomials are listed below:

1. Any polynomial $f(\lambda) \in \mathbb{C}[\lambda]$ with $\deg f(\lambda) = n$ can be written in the form
$$f(\lambda) = a_n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n) \,. \tag{1.4}$$
This representation is unique up to permutation of the factors. Some of the numbers $\lambda_1, \ldots, \lambda_n$ that are the roots of the polynomial $f(\lambda)$, could be equal. In that case the product (1.4) is represented by
$$f(\lambda) = a_n(\lambda - \lambda_1)^{\mu_1} \cdots (\lambda - \lambda_q)^{\mu_q}, \quad \mu_1 + \ldots + \mu_q = n, \tag{1.5}$$
where all λ_i, $(i = 1, \ldots, q)$ are different. The number μ_i, $(i = 1, \ldots, q)$ is called the multiplicity of the root λ_i. If $f(\lambda) \in \mathbb{R}[\lambda]$ then a_n is a real number, and in the products (1.4), (1.5) for every complex root λ_i there exists the conjugate complex root with equal multiplicity.
2. For given polynomials $f(\lambda)$, $d(\lambda) \in \mathbb{F}[\lambda]$ there exists a uniquely determined pair of polynomials $q(\lambda)$, $r(\lambda) \in \mathbb{F}[\lambda]$, such that
$$f(\lambda) = q(\lambda)d(\lambda) + r(\lambda) \,, \tag{1.6}$$
where
$$\deg r(\lambda) < \deg d(\lambda) \,.$$
Hereby, the polynomial $q(\lambda)$ is called the *entire part*, and the polynomial $r(\lambda)$ is the *remainder* from the division of $f(\lambda)$ by $d(\lambda)$.
3. Let us have $f(\lambda)$, $g(\lambda) \in \mathbb{F}[\lambda]$. It is said, that the polynomial $g(\lambda)$ is a *divisor* of $f(\lambda)$, and we write $g(\lambda)|f(\lambda)$, if
$$f(\lambda) = q(\lambda)g(\lambda)$$
is true, where $q(\lambda)$ is a certain polynomial.
The greatest common divisor (GCD) of the polynomials $f_1(\lambda)$ and $f_2(\lambda)$ should be designated by $p(\lambda)$. At the same time the GCD is a divisor of $f_1(\lambda)$ and $f_2(\lambda)$, and it possesses the greatest possible degree . Up to

equivalence, the GCD is uniquely determined. Any GCD $p(\lambda)$ permits a representation of the form

$$p(\lambda) = f_1(\lambda)m_1(\lambda) + f_2(\lambda)m_2(\lambda)\,,$$

where $m_1(\lambda)$, $m_2(\lambda)$ are certain polynomials in $\mathbb{F}[\lambda]$.

4. The two polynomials $f_1(\lambda)$ and $f_2(\lambda)$ are called *coprime* if their monic GCD is equal to one, that means, up to constants, these polynomials possess no common divisors. For the polynomials $f_1(\lambda)$ and $f_2(\lambda)$ to be coprime, it is necessary and sufficient that there exist polynomials $m_1(\lambda)$ and $m_2(\lambda)$ with
$$f_1(\lambda)m_1(\lambda) + f_2(\lambda)m_2(\lambda) = 1\,.$$
5. If
$$f_1(\lambda) = p(\lambda)\tilde{f}_1(\lambda)\,, \qquad f_2(\lambda) = p(\lambda)\tilde{f}_2(\lambda)\,,$$
where $p(\lambda)$ is a GCD of $f_1(\lambda)$ and $f_2(\lambda)$, then the polynomials $\tilde{f}_1(\lambda)$ and $\tilde{f}_2(\lambda)$ are coprime.

1.3 Matrices over Rings

1. Let $\mathcal{N}$ be a commutative ring with unit element forming an integrity region, such that $ab = 0$ implies a or b equal to zero, where 0 is the zero element of the ring $\mathcal{N}$. Then from $ab = 0$, $a \neq 0$ it always follows $b = 0$.

2. The rectangular scheme

$$A = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \tag{1.7}$$

is named a rectangular *matrix over the ring* $\mathcal{N}$, where the a_{ik} , $(i = 1,\dots,n;\ k = 1,\dots,m)$ are elements of the ring $\mathcal{N}$. In what follows, the set of matrices is designated by $\mathcal{N}_{nm}$. The integers n and m are called the *dimension* of the matrix. In case of $m = n$ we speak of a *quadratic* matrix A, for $m < n$ of a *vertical* and for $m > n$ of a *horizontal* matrix A. For matrices over rings the operations addition, (scalar) multiplication with elements of the ring $\mathcal{N}$, multiplication of matrices by matrices and transposition are defined. All these operations are defined in the same way as for matrices over numbers [51, 44].

3. Every quadratic matrix $A \in \mathcal{N}_{nn}$ is related to its determinant $\det A$ which is calculated in the same way as for number matrices. However, in the given case the value of $\det A$ is an element of the ring $\mathcal{N}$. A matrix A with $\det A \neq 0_{\mathcal{N}}$ is called *regular* or *non-singular*, for $\det A = 0_{\mathcal{N}}$ it is called *singular*.

4. For any matrix $A \in \mathcal{N}_{nn}$ there uniquely exists a matrix $\operatorname{adj} A$ of the form

$$\operatorname{adj} A = \begin{bmatrix} A_{11} & \dots & A_{n1} \\ \vdots & \vdots & \vdots \\ A_{1n} & \dots & A_{nn} \end{bmatrix}, \tag{1.8}$$

where A_{ik} is the algebraic complement (the adjoint) of the element a_{ik} of the matrix A, which is received as the determinant of those matrix that remains by cutting the $i-$th row and $k-$th column multiplied by the sign-factor $(-1)^{i+k}$. The matrix $\operatorname{adj} A$ is called the *adjoint of the matrix* A. The matrices A and $\operatorname{adj} A$ are connected by the relation

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = (\det A)I_n\,, \tag{1.9}$$

where the identity matrix I_n is defined by

$$I_n = \begin{bmatrix} 1_\mathcal{N} & 0_\mathcal{N} & \dots & 0_\mathcal{N} \\ 0_\mathcal{N} & 1_\mathcal{N} & \dots & 0_\mathcal{N} \\ \vdots & \vdots & \ddots & \vdots \\ 0_\mathcal{N} & 0_\mathcal{N} & \dots & 1_\mathcal{N} \end{bmatrix} = \operatorname{diag}\{1_\mathcal{N},\ \dots\ ,\ 1_\mathcal{N}\}$$

with the unit element $1_\mathcal{N}$ of the ring $\mathcal{N}$, and diag means the diagonal matrix.

5. In the following, matrices of dimension $n \times 1$ are called as *columns* and matrices of dimension $1 \times m$ as *rows*, and both are referred to as *vectors* . The number n is named the *height* of the column, and the number m the *width* of the row, and both are the *length* of the vector.

Let $u_1, u_2, \dots, u_k$ be rows of $\mathcal{N}_{1m}$. As a *linear combination* of the rows $u_1, \dots, u_k$, we term the row

$$\tilde{u} = c_1u_1 + \dots + c_ku_k\,,$$

where the c_i, $(i = 1, \dots, k)$ are elements of the ring $\mathcal{N}$. The set of rows $\{u_1, \dots, u_k\}$ is named *linear dependent*, if there exist coefficients $c_1, \dots, c_k$, that are not all equal to zero, such that $\tilde{u} = O_{1m}$. Here and furthermore, O_{nm} designates the *zero matrix*, *i.e.* that matrix in $\mathcal{N}_{nm}$ having all its elements equal to the zero element $0_\mathcal{N}$.

If the equation

$$cu = c_1u_1 + \dots + c_ku_k$$

is valid with a $c \neq 0_\mathcal{N}$, then we say that the column u depends linearly on the columns $u_1, \dots, u_k$.

For the set $\{u_1, \dots, u_k\}$ of columns to be linear dependent, it is necessary and sufficient that one column depends linearly on the others in the sense of the above definition.

For rows over the ring $\mathcal{N}$ the important statement is true: Any set of rows of the width m with more than m elements is linear dependent. In analogy, any set of columns of height n with more than n elements is also linear dependent.

6. Let a finite or infinite set $\mathcal{U}$ of rows of width m be given. Furthermore, let r be the maximal number of linear independent elements of $\mathcal{U}$, where due to the above statement $r \leq m$ is valid. An arbitrary subset of r linear independent rows of $\mathcal{U}$ is called a *basis* of the set $\mathcal{U}$, , the number r itself is called the *normal rank* of $\mathcal{U}$. All that is said above can be directly transferred to sets of columns.

7. Let a matrix $A \in \mathcal{N}_{nm}$ be given, and $\mathcal{U}$ should be the set of rows of A, and $\mathcal{V}$ the set of its columns. Then the following important statements take place:

1. The normal rank of the set $\mathcal{U}$ of the rows of the matrix A is equal to the normal rank of the set $\mathcal{V}$ of its columns. The common value of these ranks is called the *normal rank of the matrix* A, and it is designated by $\operatorname{rank} A$.
2. The normal rank of the matrix A is equal to the highest order of its subdeterminants (*minors*) different from zero. (Here zero again means the zero element of the ring $\mathcal{N}$.)
3. For the linear independence of all rows (columns) of a quadratic matrix, it is necessary and sufficient that it is non-singular.

For arbitrary matrices $A \in \mathcal{N}_{nm}$, the above statements imply

$$\operatorname{rank} \mathcal{V} = \operatorname{rank} \mathcal{U} \leq \min(n, m) \stackrel{\triangle}{=} \gamma_A \,.$$

Hereinafter, the symbol '$\stackrel{\triangle}{=}$' stands for equality by definition. In the following, we say that the matrix A has *maximal or full normal* rank, if $\operatorname{rank} A = \gamma_A$, or that it is *non-degenerated.* In the following the symbol 'rank' also denotes the rank of an ordinary number matrix. This notation does not lead to contradictions, because for matrices over the fields of real or complex numbers the normal rank coincides with the ordinary rank.

8. For Matrix (1.7), the expression

$$A \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = \det \begin{bmatrix} a_{i_1 k_1} & \ldots & a_{i_1 k_p} \\ \vdots & \vdots & \vdots \\ a_{i_p k_1} & \ldots & a_{i_p k_p} \end{bmatrix}$$

denotes the minor of the matrix A, which is calculated by the elements, that are at the same time members of the rows with the numbers $i_1, \ldots, i_p$, and of the columns with the numbers $k_1, \ldots, k_p$. Let

$$C = AB$$

be given with $C \in \mathcal{N}_{nm}$, $A \in \mathcal{N}_{n\ell}$, $B \in \mathcal{N}_{\ell m}$. Then if $n = m$, the matrix C is quadratic, and for $n \leq \ell$ we have

$$\det C = \sum_{1\le k_1<\cdots<k_n\le \ell} \cdots \sum A\begin{pmatrix} 1 & 2 & \dots & n \\ k_1 & k_2 & \dots & k_n \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

This relation is called *Binet-Cauchy-formula* [51]. For $n > \ell$ we obtain $\det C = 0$.

The formula of Binet-Cauchy permits to express an arbitrary minor of a product by the corresponding minors of its factors. For $p \le \ell$ this formula takes the form

$$C\begin{pmatrix} i_1 & \dots & i_p \\ j_1 & \dots & j_p \end{pmatrix} = \sum_{1\le k_1<\cdots<k_p\le \ell} \cdots \sum A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}.$$

For $p > \ell$, all minors in this relation will be equal to zero.

1.4 Polynomial Matrices

1. By a *polynomial matrix* $A(\lambda)$ we mean a matrix of the form

$$A(\lambda) = \begin{bmatrix} a_{11}(\lambda) & \dots & a_{1m}(\lambda) \\ \vdots & \vdots & \vdots \\ a_{n1}(\lambda) & \dots & a_{nm}(\lambda) \end{bmatrix},$$

where all elements are polynomials in $\mathbb{F}[\lambda]$, especially also in $\mathbb{R}[\lambda]$ or $\mathbb{C}[\lambda]$. The set of these matrices will be designated by $\mathbb{F}_{nm}[\lambda]$, or symbolised directly by $\mathbb{R}_{nm}[\lambda]$ resp. $\mathbb{C}[\lambda]$, and their subsets containing the constant matrices, are denoted by $\mathbb{F}_{nm}$, $\mathbb{R}_{nm}$ or $\mathbb{C}_{nm}$, respectively. The matrices in $\mathbb{R}_{nm}$ and $\mathbb{R}_{nm}[\lambda]$ are called real.

2. Let, especially

$$u_i(\lambda) = \begin{bmatrix} a_{i1}(\lambda) & \dots & a_{im}(\lambda) \end{bmatrix}, \quad (i = 1, 2, \dots, p)$$

be a certain set of rows with width m. The above defined rows will be called *linear dependent* in $\mathbb{F}_{1m}[\lambda]$, if and only if there exist polynomials $c_i(\lambda) \in \mathbb{F}[\lambda]$ that are not all zero at the same time, such that

$$\sum_{i=1}^{p} c_i(\lambda) u_i(\lambda) = O_{1m}.$$

Here $O_{\ell m}$ is the matrix of dimension $\ell \times m$ with all elements equal to the zero polynomial.

Based on this fundamental understanding of these definitions, all derived concepts and insight can be transferred to polynomial matrices, especially the normal rank of matrices over rings and also the formula of Binet-Cauchy.

3. Any polynomial matrix $A(\lambda) \in \mathbb{F}_{nm}[\lambda]$ can be written in the form

$$A(\lambda) = A_0\lambda^q + A_1\lambda^{q-1} + \ldots + A_q\,, \tag{1.10}$$

where A_i, $(i = 1, \ldots, q)$ are constant matrices in $\mathbb{F}_{nm}$. The matrix A_0 is named the *highest coefficient* of the polynomial matrix $A(\lambda)$. If $A_0 \neq O_{nm}$ is true, then the number q is called the *degree of the polynomial matrix* $A(\lambda)$, and it is designated by $\deg A(\lambda)$. If $n \neq m$, or $\det A(\lambda) \equiv 0$ in case of $n = m$, the matrix $A(\lambda)$ is called *singular*. For $\det A(\lambda) \not\equiv 0$ the matrix $A(\lambda)$ is called *non-singular*. A non-singular matrix (1.10) is called *regular*, if $\det A_0 \neq 0$, and *anomalous*, if $\det A_0 = 0$ is true.

In the general case we have

$$\deg[A(\lambda)B(\lambda)] \leq \deg A(\lambda) + \deg B(\lambda)\,. \tag{1.11}$$

However, if one of the factors is regular, then

$$\deg[A(\lambda)B(\lambda)] = \deg A(\lambda) + \deg B(\lambda)\,. \tag{1.12}$$

4. For $A(\lambda) \in \mathbb{F}_{nn}[\lambda]$, Matrix (1.10) is related to its determinant $\det A(\lambda)$, which itself is a polynomial in $\mathbb{F}[\lambda]$. In accordance with the above statements the matrix is non-singular if its determinant is different from the zero polynomial. A non-singular matrix $A(\lambda)$ is related to the non-negative number

$$\operatorname{ord} A(\lambda) = \deg \det A(\lambda)$$

that is called the *order* of the matrix $A(\lambda)$. The degree and order of a matrix $A(\lambda)$ are connected by the inequalities

$$\operatorname{ord} A(\lambda) \leq n \deg A(\lambda) \tag{1.13}$$

or

$$\deg A(\lambda) \geq \frac{1}{n} \operatorname{ord} A(\lambda)\,. \tag{1.14}$$

For a regular matrix $A(\lambda)$ the inequalities (1.13), (1.14) become equalities. In general for a given order $\operatorname{ord} A(\lambda)$, the degree of a matrix $A(\lambda)$ can be an arbitrary large number. A non-singular quadratic polynomial matrix $A(\lambda)$ with $\operatorname{ord} A(\lambda) = 0$, *i.e.* $\det A(\lambda) = \text{const.} \neq 0$ is called *unimodular*.

Example 1.1. The matrix

$$A(\lambda) = \begin{bmatrix} \lambda - 2 & 1 \\ \lambda^4 - 5\lambda^3 + 6\lambda^2 - 5\lambda + 6 & \lambda^3 - 3\lambda^2 + 4\lambda - 4 \end{bmatrix}$$

can be written in the form (1.10) as

$$A(\lambda) = A_0\lambda^4 + A_1\lambda^3 + A_2\lambda^2 + A_3\lambda + A_4\,,$$

where

$$A_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ -5 & 1 \end{bmatrix},$$
$$A_2 = \begin{bmatrix} 0 & 0 \\ 6 & -3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 \\ -5 & 4 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -2 & 1 \\ 6 & -4 \end{bmatrix}.$$

In the present case we have $\deg A(\lambda) = 4$. The matrix $A(\lambda)$ is non-singular, because of $n = m = 2$ and $\det A(\lambda) \not\equiv 0$. At the same time $\operatorname{ord} A(\lambda) = 2$ due to

$$\det A(\lambda) = 4\lambda^2 - 7\lambda + 2 .$$

Moreover, the matrix $A(\lambda)$ is anomalous, because $\det A_0 = 0$.

□

Example 1.2. Let

$$A(\lambda) = \begin{bmatrix} 1 & \lambda^2 + 4 \\ \lambda^3 - 3\lambda + 5 & \lambda^5 + \lambda^3 + 5\lambda^2 - 12\lambda + 21 \end{bmatrix}$$

be given. In this case we have $\deg A(\lambda) = 5$. At the same time $\det A(\lambda) = 1$ and $\operatorname{ord} A(\lambda) = 0$, thus the matrix $A(\lambda)$ is unimodular. □

1.5 Left and Right Equivalence of Polynomial Matrices

1. Let two polynomial matrices $A_1(\lambda)$, $A_2(\lambda) \in \mathbb{F}_{nm}[\lambda]$ be given. The matrices $A_1(\lambda)$ and $A_2(\lambda)$ are called *left-equivalent*, if one of them can be generated from the other by applying the following operations, which are called *left elementary operations*:

1. Exchange of two rows
2. Multiplying the elements of any row with one and the same non-zero number in $\mathbb{F}$
3. Adding to the elements of any row, the by one and the same polynomial in $\mathbb{F}[\lambda]$ multiplied corresponding elements of another row.

It is known that matrices $A_1(\lambda)$, $A_2(\lambda)$ are left-equivalent if and only if there exists a unimodular matrix $p(\lambda)$, such that

$$A_1(\lambda) = p(\lambda)A_2(\lambda) .$$

2. By applying left elementary operations, any matrix $A(\lambda)$ can be given a special form that later on is named the left canonical form of Hermite.

Theorem 1.3 (following [113]). *Let the matrix $A(\lambda) \in \mathbb{F}_{nm}[\lambda]$ have maximal rank γ_A, and the first γ_A columns of $A(\lambda)$ should have a minor of non-vanishing order γ_A. Then in dependence of its dimension, the matrix $A(\lambda)$ can be transformed by left elementary operations into one of the three forms:*

$\boldsymbol{\gamma_A = m = n:}$

$$p(\lambda)A(\lambda) = \begin{bmatrix} g_{11}(\lambda) & g_{12}(\lambda) & \dots & g_{1n}(\lambda) \\ 0 & g_{22}(\lambda) & \dots & g_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g_{nn}(\lambda) \end{bmatrix} \triangleq \tilde{A}_l(\lambda)\,, \tag{1.15}$$

$\boldsymbol{\gamma_A = n < m:}$

$$p(\lambda)A(\lambda) = \begin{bmatrix} g_{11}(\lambda) & g_{12}(\lambda) & \dots & g_{1n}(\lambda) & g_{1,n+1}(\lambda) & \dots & g_{1m}(\lambda) \\ 0 & g_{22}(\lambda) & \dots & g_{2n}(\lambda) & g_{2,n+1}(\lambda) & \dots & g_{2m}(\lambda) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & g_{nn}(\lambda) & g_{n,n+1}(\lambda) & \dots & g_{nm}(\lambda) \end{bmatrix} \triangleq \tilde{A}_l(\lambda)\,, \tag{1.16}$$

$\boldsymbol{\gamma_A = m < n:}$

$$p(\lambda)A(\lambda) = \begin{bmatrix} g_{11}(\lambda) & g_{12}(\lambda) & \dots & g_{1m}(\lambda) \\ 0 & g_{22}(\lambda) & \dots & g_{2m}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g_{mm}(\lambda) \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \triangleq \tilde{A}_l(\lambda)\,. \tag{1.17}$$

In (1.15)–(1.17) the matrix $p(\lambda)$ is unimodular, and the $g_{ii}(\lambda)$ are monic polynomials, where every $g_{ii}(\lambda)$ is of highest degree in its column. Doing so, the matrix $\tilde{A}_l(\lambda)$ is uniquely determined by $A(\lambda)$. Moreover, in Formulae (1.15) and (1.16) the matrix $p(\lambda)$ is also uniquely committed. ■

In the following, the suitable matrix $\tilde{A}_l(\lambda)$ is said to be the *left canonical form* of the corresponding matrix $A(\lambda)$ or also its *left Hermitian form.*

3. By *right elementary operations* , we understand the above declared operations for columns instead of rows. Two matrices are called *right-equivalent*, if we can generate any one from the other by applying right elementary operations. Two polynomial matrices $A_1(\lambda)$, $A_2(\lambda)$ are right-equivalent if and only if there exists a unimodular matrix $q(\lambda)$ with

$$A_1(\lambda) = A_2(\lambda)q(\lambda)\,.$$

In analogy to Theorem 1.3 the following theorem holds.

Theorem 1.4. *Let the matrix $A(\lambda)$ have the maximal rank γ_A, and the first γ_A rows of $A(\lambda)$ should possess a non-zero minor of order γ_A. Then according to its dimension, by applying right elementary operations, the matrix $A(\lambda)$ can be transformed into one of the three forms:*

$\boldsymbol{\gamma_A = m = n:}$

$$A(\lambda)q(\lambda) = \begin{bmatrix} g_{11}(\lambda) & 0 & \dots & 0 \\ g_{21}(\lambda) & g_{22}(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1}(\lambda) & g_{n2}(\lambda) & \dots & g_{nn}(\lambda) \end{bmatrix} \triangleq \tilde{A}_r(\lambda)\,, \tag{1.18}$$

$\boldsymbol{\gamma_A = n < m:}$

$$A(\lambda)q(\lambda) = \begin{bmatrix} g_{11}(\lambda) & 0 & & \dots & 0\dots 0 \\ g_{21}(\lambda) & g_{22}(\lambda) & 0 & \dots & 0\dots 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \dots \vdots \\ g_{n1}(\lambda) & g_{n2}(\lambda) & \dots & g_{nn}(\lambda) & 0\dots 0 \end{bmatrix} \triangleq \tilde{A}_r(\lambda)\,, \tag{1.19}$$

$\boldsymbol{\gamma_A = m < n:}$

$$A(\lambda)q(\lambda) = \begin{bmatrix} g_{11}(\lambda) & 0 & \dots & 0 \\ g_{21}(\lambda) & g_{22}(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1}(\lambda) & g_{m2}(\lambda) & \dots & g_{mm}(\lambda) \\ g_{m+1,1}(\lambda) & g_{m+1,2}(\lambda) & \dots & g_{m+1,m}(\lambda) \\ \vdots & \vdots & \dots & \vdots \\ g_{n1}(\lambda) & g_{n2}(\lambda) & \dots & g_{nm}(\lambda) \end{bmatrix} \triangleq \tilde{A}_r(\lambda)\,. \tag{1.20}$$

In (1.18)–(1.20) the matrix $q(\lambda)$ is unimodular, and the $g_{ii}(\lambda)$ are monic polynomials, where every $g_{ii}(\lambda)$ has the highest degree in its row. Doing so, the matrix $\tilde{A}_r(\lambda)$ is uniquely determined by $A(\lambda)$. Moreover, the matrix $q(\lambda)$ in (1.18) and (1.20) is also uniquely committed. ∎

The suitable matrix $\tilde{A}_r(\lambda)$ in (1.18)–(1.20) is said to be the *right canonical form* of the polynomial matrix $A(\lambda)$, or its *right Hermitian form.*

Example 1.5. Let

$$A(\lambda) = \begin{bmatrix} \lambda^4 + 1 & \lambda^4 + 3 \\ \lambda^6 + 2\lambda^2 + 1 & \lambda^6 + 4\lambda^2 + 2 \end{bmatrix}.$$

In this case we have $\det A(\lambda) = \lambda^4 - 2\lambda^2 - 1$. Hence $\deg A(\lambda) = 6$ and $\operatorname{ord} A(\lambda) = 4$. The matrix

$$p(\lambda) = \begin{bmatrix} 0.5 & 0.5(1-\lambda^2) \\ -(\lambda^2+1) & \lambda^4 + 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\lambda^2 & 1 \end{bmatrix}$$

is unimodular. By direct calculation we confirm

$$p(\lambda)A(\lambda) = \begin{bmatrix} 1 & 2.5 - 0.5\lambda^2 \\ 0 & \lambda^4 - 2\lambda^2 - 1 \end{bmatrix},$$

which has degree 4. The matrix on the right side of the last equation is a Hermitian canonical form. As a conclusion of Theorem 1.3, it follows that this Hermitian form $\tilde{A}_l(\lambda)$ and its transformation matrix $p(\lambda)$ are uniquely determined. It should be remarked that the matrix $\tilde{A}_l(\lambda)$ possesses not the smallest possible degree of all matrices that are left-equivalent to $A(\lambda)$. Indeed, consider the product

$$A_1(\lambda) = \begin{bmatrix} 1 & -\lambda^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\lambda^2 & 1 \end{bmatrix} A(\lambda) = \begin{bmatrix} 1 - \lambda^2 & 3 - \lambda^2 \\ \lambda^2 + 1 & \lambda^2 + 2 \end{bmatrix}.$$

Then obviously $\deg A_1(\lambda) = 2$. This is the minimal degree, and this result confirms Inequality (1.14). □

1.6 Row and Column Reduced Matrices

1. Let the non-singular quadratic matrix $A(\lambda) \in \mathbb{F}_{nn}[\lambda]$ be given, and $a_1(\lambda), \ldots, a_n(\lambda)$ be the rows of $A(\lambda)$. With the notation

$$\alpha_i = \deg a_i(\lambda)\,, \quad (i = 1, \ldots, n)\,,$$

the matrix $A(\lambda)$ can be written in the form

$$A(\lambda) = \operatorname{diag}\{\lambda^{\alpha_1}, \ldots, \lambda^{\alpha_n}\}A_0 + A_1(\lambda)\,. \tag{1.21}$$

Herein $A_1(\lambda)$ is a matrix, where the degree of its i-th row is smaller than α_i, and A_0 is a constant matrix. Formula (1.21) could be transformed into

$$A(\lambda) = \operatorname{diag}\{\lambda^{\alpha_1}, \ldots, \lambda^{\alpha_n}\}\left(A_0 + A_1\lambda^{-1} + \ldots + A_p\lambda^{-p}\right)\,, \tag{1.22}$$

where $p \geq 0$ is an integer, and the A_i are constant matrices. The number

$$\alpha_l = \alpha_1 + \ldots + \alpha_n$$

is named the *left order* of the matrix $A(\lambda)$. Denote

$$\alpha_{\max} = \max_{1\leq i\leq n}\{\alpha_i\}\,.$$

Then obviously

$$\deg A(\lambda) = \alpha_{\max}\,. \tag{1.23}$$

In analogy, assuming that $b_1(\lambda), \ldots, b_n(\lambda)$ are the columns of $A(\lambda)$ and

$$\beta_i = \deg b_i(\lambda)\,,$$

we generate the representation

$$A(\lambda) = \left(B_0 + B_1\lambda^{-1} + \ldots + B_q\lambda^{-q}\right) \operatorname{diag}\{\lambda^{\beta_1}, \ldots, \lambda^{\beta_n}\}\,. \tag{1.24}$$

The number

$$\beta_r = \beta_1 + \ldots + \beta_n$$

is called the *right order* of the matrix $A(\lambda)$. Introduce the notation

$$\beta_{\max} = \max_{1\le i\le n} \{\beta_i\}\,.$$

Then we obtain

$$\deg A(\lambda) = \beta_{\max}\,.$$

Example 1.6. [68]: Consider the matrix

$$A(\lambda) = \begin{bmatrix} \lambda^2 - 1 & \lambda & -3\lambda \\ \lambda^2 & \lambda - 1 & 2\lambda - 1 \\ \lambda + 2 & -\lambda & 2 \end{bmatrix}. \tag{1.25}$$

In this case we have $\alpha_1 = 2$, $\alpha_2 = 2$, $\alpha_3 = 1$, and the left order of the matrix $A(\lambda)$ becomes $\alpha_l = 5$. In the representation (1.21) we get

$$A_0 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \quad A_1(\lambda) = \begin{bmatrix} -1 & \lambda & -3\lambda \\ 0 & \lambda - 1 & 2\lambda - 1 \\ 2 & 0 & 2 \end{bmatrix}$$

and therefore, (1.22) yields

$$A(\lambda) = \operatorname{diag}\{\lambda^2,\ \lambda^2,\ \lambda\}\left(A_0 + A_1\lambda^{-1} + A_2\lambda^{-2}\right) \tag{1.26}$$

with

$$A_0 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 1 & -3 \\ 0 & 1 & 2 \\ 2 & 0 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{1.27}$$

At the same time we have $\beta_1 = 2$, $\beta_2 = 1$, $\beta_3 = 1$, and the right order of $A(\lambda)$ becomes $\beta_r = 4$. The representation (1.24) takes the form

$$A(\lambda) = \left(B_0 + B_1\lambda^{-1} + B_2\lambda^{-2}\right) \operatorname{diag}\{\lambda^2,\ \lambda,\ \lambda\}\,,$$

where

$$B_0 = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 1 & 2 \\ 0 & -1 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 2 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}.$$

$\square$

2. The matrix $A(\lambda)$ is said to be *row reduced*, if we have in the representation (1.21)

$$\det A_0 \neq 0 \tag{1.28}$$

and it said to be *column reduced*, if in the representation (1.24)

$$\det B_0 \neq 0$$

is true.

Column-reduced matrices can be generated from row-reduced matrices simply by transposition. Therefore, in the following only row-reduced matrices will be considered.

Lemma 1.7. *For Matrix (1.21) to be row reduced, a necessary and sufficient condition is the validity of the equation*

$$\operatorname{ord} A(\lambda) = \deg \det A(\lambda) = \alpha_l \,. \tag{1.29}$$

Proof. From (1.21) we get

$$\det A(\lambda) = \lambda^{\alpha_l} \det A_0 + a_1(\lambda)$$

with $\deg a_1(\lambda) < \alpha_l$. For (1.29) to be valid, (1.28) is necessary. If, conversely, (1.29) is fulfilled, then $\det A_0 \neq 0$ is true, and the matrix $A(\lambda)$ is row reduced. ∎

Example 1.8. For Matrix (1.25) we get $\det A_0 = 0$, $\det B_0 = 5$, therefore the matrix $A(\lambda)$ is column reduced but not row reduced. Hereby, we obtain $\operatorname{ord} A(\lambda) = \beta_r = 4$. □

Theorem 1.9 ([133]). *Any non-singular matrix $A(\lambda)$ can be made row-reduced by left-equivalent transforms.*

Proof. Assume the matrix $A(\lambda)$ be given in form of Representation (1.22). Then for $\det A_0 \neq 0$ the matrix $A(\lambda)$ is already row reduced. Therefore, take a singular matrix A_0, *i.e.* $\det A_0 = 0$. Then there exists a non-zero row vector $\nu = (\nu_1, \ldots, \nu_n)$ such that

$$\nu A_0 = O_{1n} \tag{1.30}$$

is fulfilled. Let $\nu_{i_1}, \ldots, \nu_{i_q}$, $(1 \leq q \leq n)$ be the non-zero components of ν, and $\alpha_{i_1}, \ldots, \alpha_{i_q}$ are the corresponding exponents α_i. Denote

$$\psi = \max_{1 \leq j \leq q} \{\alpha_{i_j}\} \,,$$

and let γ be a value of the index j, for which $\alpha_{i_\gamma} = \psi$ is valid. Then the row

$$\nu(\lambda) = \left[\, \nu_1 \lambda^{\psi - \alpha_1} \;\; \nu_2 \lambda^{\psi - \alpha_2} \;\; \ldots \;\; \nu_\gamma \;\; \ldots \;\; \nu_{n-1} \lambda^{\psi - \alpha_{n-1}} \;\; \nu_n \lambda^{\psi - \alpha_n} \,\right] \tag{1.31}$$

comes out as a polynomial. Now consider the matrix $P(\lambda)$ that is generated from the identity matrix I_n by exchanging the γ-th row by the row (1.31):

$$P(\lambda)=\begin{bmatrix} 1 & 0 & \dots\ 0\ \dots & 0 & 0 \\ \vdots & \ddots & \dots\ \vdots\ \dots & \vdots & \vdots \\ \nu_1\lambda^{\psi-\alpha_1} & \nu_2\lambda^{\psi-\alpha_2} & \dots\ \nu_\gamma\ \dots & \nu_{n-1}\lambda^{\psi-\alpha_{n-1}} & \nu_n\lambda^{\psi-\alpha_n} \\ \vdots & \vdots & \dots\ \vdots\ \dots & \ddots & \vdots \\ 0 & 0 & \dots\ 0\ \dots & 0 & 1 \end{bmatrix}. \tag{1.32}$$

Due to $\det P(\lambda)=\nu_\gamma\neq 0$, the matrix $P(\lambda)$ is unimodular. Therefore, as shown in [133], the equation

$$P(\lambda)A(\lambda)=\operatorname{diag}\{\lambda^{\alpha_1},\dots,\lambda^{\alpha_n}\}\left(\tilde{A}_0+\tilde{A}_1\lambda^{-1}+\dots+\tilde{A}_p\lambda^{-p}\right) \tag{1.33}$$

holds with

$$\tilde{A}_i=DA_i\,, \tag{1.34}$$

and the matrix D is generated from the identity matrix I_n by exchanging the γ-th row by the row ν:

$$D=\begin{bmatrix} 1 & 0 & \dots\ 0\ \dots & 0 & 0 \\ \vdots & \ddots & \dots\ \vdots\ \dots & \vdots & \vdots \\ \nu_1 & \nu_2 & \dots\ \nu_\gamma\ \dots & \nu_{n-1} & \nu_n \\ \vdots & \vdots & \dots\ \vdots\ \dots & \ddots & \vdots \\ 0 & 0 & \dots\ 0\ \dots & 0 & 1 \end{bmatrix}. \tag{1.35}$$

Obviously, $\det D=\nu_\gamma\neq 0$. From (1.30) and (1.33) it follows that the γ-th row of the matrix $\tilde{A}_0$ is identical to zero. That means, Equation (1.33) can be written in the form

$$\begin{aligned} P(\lambda)A(\lambda)=\operatorname{diag}\{\lambda^{\alpha_1},\dots,\lambda^{\alpha_{\gamma-1}},\lambda^{\alpha_\gamma-1},\lambda^{\alpha_{\gamma+1}},\dots,\lambda^{\alpha_n}\}\,\cdot \\ \left(\tilde{\tilde{A}}_0+\tilde{\tilde{A}}_1\lambda^{-1}+\dots+\tilde{\tilde{A}}_p\lambda^{-p}\right)\,, \end{aligned} \tag{1.36}$$

where the matrices $\tilde{\tilde{A}}_i$ $(i=0,\dots,p-1)$ are built from the matrices $\tilde{A}_i$ by substituting their γ-th rows by the γ-th row of $\tilde{A}_{i+1}$. Hereby, the γ-th row of $\tilde{A}_p$ is substituted by the zero row. If in Relation (1.36) the matrix $\tilde{\tilde{A}}_0$ is regular, then the matrix $P(\lambda)A(\lambda)$ is row reduced, and the transformation procedure finishes. If, however, the matrix $\tilde{\tilde{A}}_0$ is already singular, we have to repeat the transformation procedure again. It was shown in [133] that for a non-singular matrix $A(\lambda)$ this algorithm after a finite number of steps yields a row-reduced matrix. ■

Example 1.10. Generate a row-reduced form of the matrix $A(\lambda)$ in (1.26), (1.27). Equation (1.30) leads to the system of linear equations

$$\nu_1 + \nu_2 + \nu_3 = 0\,, \qquad -\nu_3 = 0\,,$$

so we choose $\nu = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}$, $\gamma = 1$ and $\psi = 2$. Applying (1.31), (1.32) and (1.34) yields

$$P(\lambda) = D = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Using this result and (1.34), we find

$$\tilde{A}_0 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \quad \tilde{A}_1 = \begin{bmatrix} 0 & 0 & -5 \\ 0 & 1 & 2 \\ 2 & 0 & 2 \end{bmatrix}, \quad \tilde{A}_2 = \begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now, exchange the first row of $\tilde{A}_0$ by the first row of $\tilde{A}_1$, the first row of $\tilde{A}_1$ by the first row of $\tilde{A}_2$, and the first row of $\tilde{A}_2$ by the zero row. As result we get

$$\tilde{\tilde{A}}_0 = \begin{bmatrix} 0 & 0 & -5 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}, \quad \tilde{\tilde{A}}_1 = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 2 \end{bmatrix}, \quad \tilde{\tilde{A}}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrix $\tilde{\tilde{A}}_0$ is regular. Therefore, the procedure stops, and with the help of (1.36) we get

$$\begin{aligned} P(\lambda)A(\lambda) &= \operatorname{diag}\{\lambda, \lambda^2, \lambda\} \left(\tilde{\tilde{A}}_0 + \tilde{\tilde{A}}_1\lambda^{-1} + \tilde{\tilde{A}}_2\lambda^{-2}\right) \\ &= \begin{bmatrix} -1 & 1 & -5\lambda + 1 \\ \lambda^2 & \lambda - 1 & 2\lambda - 1 \\ \lambda + 2 & -\lambda & 2 \end{bmatrix}. \end{aligned}$$

□

3. A number of useful properties of row-reduced matrices follows from the above explanations.

Theorem 1.11. (see [69]) *Let the matrices*

$$\begin{aligned} A(\lambda) &= \operatorname{diag}\{\lambda^{\alpha_1}, \ldots, \lambda^{\alpha_n}\} \left(\tilde{A}_0 + \tilde{A}_1\lambda^{-1} + \ldots\right), \\ B(\lambda) &= \operatorname{diag}\{\lambda^{\phi_1}, \ldots, \lambda^{\phi_n}\} \left(\tilde{B}_0 + \tilde{B}_1\lambda^{-1} + \ldots\right) \end{aligned} \tag{1.37}$$

be given. Then, if the matrices $A(\lambda)$ and $B(\lambda)$ are row reduced and left equivalent, then the sets of numbers $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\phi_1, \ldots, \phi_n\}$ coincide. ■

Corollary 1.12. *If the matrices $A(\lambda)$ and $B(\lambda)$ are left equivalent, and the matrix $A(\lambda)$ is row reduced, then*

$$\sum_{i=1}^{n} \alpha_i \leq \sum_{i=1}^{n} \phi_i$$

is true, where the equality takes place if and only if the matrix $B(\lambda)$ is also row reduced.

Proof. Because the matrices (1.37) are left equivalent, they possess the same order. Therefore, by Lemma 1.7 it follows

$$\phi_1 + \ldots + \phi_n \geq \operatorname{ord} B(\lambda) = \operatorname{ord} A(\lambda) = \alpha_1 + \ldots + \alpha_n \,,$$

where the equality in the left part exactly takes place, when the matrix $B(\lambda)$ is row reduced. ∎

Corollary 1.13. *Under the conditions of Theorem 1.11,*

$$\deg A(\lambda) = \deg B(\lambda) \,. \tag{1.38}$$

Proof. From (1.23) and (1.37), we get

$$\deg A(\lambda) = \max_{1\leq i\leq n} \{\alpha_i\} = \max_{1\leq i\leq n} \{\phi_i\} = \deg B(\lambda) \,,$$

because the sets of numbers α_i and ϕ_i coincide. ∎

Corollary 1.14. *Let the matrices $A(\lambda)$ and $B(\lambda)$ in (1.37) be left equivalent, where the matrix $A(\lambda)$ is row reduced, but the matrix $B(\lambda)$ is not row reduced. Then we have*

$$\deg A(\lambda) \leq \deg B(\lambda) \,. \tag{1.39}$$

Proof. In contrary to the claim, we assume

$$\chi = \deg B(\lambda) < \deg A(\lambda) = \alpha_{\max} \,.$$

Then $B(\lambda)$ is given in a row-reduced form by applying Relation (1.36). In this manner, we get

$$Q(\lambda)B(\lambda) = \operatorname{diag}\{\lambda^{\tilde{\phi}_1}, \ldots \lambda^{\tilde{\phi}_n}\} \left(\tilde{B}_0 + \tilde{B}_1\lambda^{-1} + \ldots\right)$$

with a unimodular matrix $Q(\lambda)$ and $\det \tilde{B}_0 \neq 0$. From (1.36), it is seen that

$$\deg[Q(\lambda)B(\lambda)] \leq \deg B(\lambda) = \chi < \alpha_{\max} = \deg A(\lambda) \,,$$

which contradicts Equation (1.38). Hence it follows (1.39). ∎

1.7 Equivalence of Polynomial Matrices

1. The matrices $A_1(\lambda)$, $A_2(\lambda) \in \mathbb{F}_{nm}[\lambda]$ are called *equivalent*, if

$$A_1(\lambda) = p(\lambda)A_2(\lambda)q(\lambda) \tag{1.40}$$

is true with unimodular matrices $p(\lambda)$, $q(\lambda)$. Obviously, left-equivalent or right-equivalent matrices are also equivalent. Formula (1.40) says that the matrices $A_1(\lambda)$ and $A_2(\lambda)$ are equivalent if and only if they could be generated of each other by left or right elementary operations.

2.

Theorem 1.15 ([51]). *Any $n \times m$ matrix $A(\lambda)$ with the normal rank ρ is equivalent to the matrix*

$$S_A(\lambda) = \begin{bmatrix} S_\rho(\lambda) & O_{\rho,m-\rho} \\ O_{n-\rho,\rho} & O_{n-\rho,m-\rho} \end{bmatrix}, \tag{1.41}$$

where the matrix $S_\rho(\lambda)$ has the form

$$S_\rho(\lambda) = \operatorname{diag}\{a_1(\lambda), \ldots, a_\rho(\lambda)\}, \tag{1.42}$$

and the $a_i(\lambda)$ are monic polynomials, where every polynomial $a_{i+1}(\lambda)$ is divisible by $a_i(\lambda)$. ∎

Matrix (1.41) is uniquely determined by the matrix $A(\lambda)$, and it is named as the *Smith-canonical form* of the matrix $A(\lambda)$.

Corollary 1.16. *It follows immediately from Relations (1.40)–(1.42), that under the condition $\operatorname{rank} A(\lambda) = \rho$, there exists only a finite set of numbers $\tilde{\lambda}_i$, $(i = 1, \ldots, q)$ such that the number matrix $A(\tilde{\lambda}_i)$ satisfies the inequality $\operatorname{rank} A(\tilde{\lambda}_i) < \rho$.* ∎

Corollary 1.17. *Let a finite number of matrices $A_1(\lambda), \ldots, A_p(\lambda)$ with $\operatorname{rank} A_i(\lambda) = \rho_i$, $(i = 1, \ldots, p)$ be given. Then for all fixed values $\tilde{\lambda}$, excluding a certain finite set, the condition $\operatorname{rank} A_i(\tilde{\lambda}) = \rho_i$ is fulfilled.* ∎

1.8 Normal Rank of Polynomial Matrices

1. Utilising the results from the preceding section, we are able to transfer known results over the rank of number matrices to the normal rank of polynomial matrices.

Theorem 1.18. *Assume*

$$D(\lambda) = A(\lambda)B(\lambda)$$

with polynomial matrices $A(\lambda)$, $B(\lambda)$, $D(\lambda)$ of sizes $n \times \ell$, $\ell \times m$ and $n \times m$, respectively. Then the relations

$$\operatorname{rank} D(\lambda) \leq \min\{\operatorname{rank} A(\lambda), \operatorname{rank} B(\lambda)\} \tag{1.43}$$

and

$$\operatorname{rank} D(\lambda) \geq \operatorname{rank} A(\lambda) + \operatorname{rank} B(\lambda) - \ell \tag{1.44}$$

are true.

Relations (1.43), (1.44) are named *inequalities of Sylvester.*

Proof. Assume

$$\operatorname{rank} A(\lambda) = \rho_A, \ \operatorname{rank} B(\lambda) = \rho_B, \ \operatorname{rank} D(\lambda) = \rho_D$$

with

$$\rho_D > \min\{\rho_A, \rho_B\}\,. \tag{1.45}$$

Then due to Corollary 1.17, there exists a value $\lambda = \tilde{\lambda}$ with

$$\operatorname{rank} A(\tilde{\lambda}) = \rho_A, \quad \operatorname{rank} B(\tilde{\lambda}) = \rho_B, \quad \operatorname{rank} D(\tilde{\lambda}) = \rho_D\,.$$

It is possible to apply the Sylvester inequalities to the number matrices

$$D(\tilde{\lambda}) = A(\tilde{\lambda})B(\tilde{\lambda})\,,$$

which gives

$$\operatorname{rank} D(\tilde{\lambda}) \leq \min\{\operatorname{rank} A(\tilde{\lambda}) \operatorname{rank} B(\tilde{\lambda})\}\,.$$

But this contradicts (1.45). This contradiction proves the validity of Inequality (1.43).

Inequality (1.44) could be proved analogously because a corresponding inequality holds for constant matrices. ∎

2. From the inequalities of Sylvester (1.43), (1.44) ensue the following relations

$$\operatorname{rank}[A(\lambda)B(\lambda)] = \operatorname{rank} B(\lambda) \quad \text{for} \ \operatorname{rank} A(\lambda) = \ell \tag{1.46}$$

and

$$\operatorname{rank}[A(\lambda)B(\lambda)] = \operatorname{rank} A(\lambda) \quad \text{for} \ \operatorname{rank} B(\lambda) = \ell\,. \tag{1.47}$$

Herein, $\operatorname{rank} A(\lambda) = \ell$ can only be fulfilled for $n \geq \ell$, and $\operatorname{rank} B(\lambda) = \ell$ only for $m \geq \ell$. Especially, Equation (1.46) is valid if the matrix $A(\lambda)$ is non-singular, and Equation (1.47) holds if matrix $B(\lambda)$ is non-singular.

3. For arbitrary matrices $A(\lambda) \in \mathbb{F}_{nm}[\lambda]$ we introduce the notation

$$\operatorname{def} A(\lambda) \stackrel{\triangle}{=} \min\{n, m\} - \operatorname{rank} A(\lambda) = \gamma_A - \operatorname{rank} A(\lambda)$$

and call it as the *normal defect* of $A(\lambda)$. Obviously, we always have $\operatorname{def} A(\lambda) \geq 0$.

Arising from this fact and the above considerations, we conclude that the rank of any matrix does not decrease if it is multiplied from the left by a vertical or square matrix with defect zero. Analogously, the multiplication from right by an horizontal or square matrix with defect zero also would not change the rank.

4. Applying the above thoughts used in the proofs of Theorem 1.18, the known statements for number matrices [51] can also be proved for polynomial matrices.

Theorem 1.19. *The matrices $A(\lambda)$, $B(\lambda)$ and the matrix $D(\lambda)$ should be connected by*

$$D(\lambda) = \begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix} .$$

Then,

$$\operatorname{rank} D(\lambda) \le \operatorname{rank} A(\lambda) + \operatorname{rank} B(\lambda) . \tag{1.48}$$

■

Theorem 1.20. *For any polynomial matrices $A(\lambda)$ and $B(\lambda)$ of equal dimension,*

$$\operatorname{rank}[A(\lambda) + B(\lambda)] \le \operatorname{rank} A(\lambda) + \operatorname{rank} B(\lambda) .$$

■

Remark 1.21. A corresponding relation to the last one was proven for number matrices in [147].

1.9 Invariant Polynomials and Elementary Divisors

1. Applying (1.41) and (1.42), the Smith-canonical form of a polynomial matrix $A(\lambda)$ can be written in the form

$$S_A(\lambda) = \left[\begin{array}{cccc|c} h_1(\lambda) & 0 & \dots & 0 & \\ 0 & h_1(\lambda)h_2(\lambda) & \dots & 0 & \\ \vdots & \vdots & \ddots & \vdots & O_{\rho,m-\rho} \\ 0 & 0 & \dots & h_1(\lambda)h_2(\lambda)\cdots h_\rho(\lambda) & \\ \hline & & O_{n-\rho,\rho} & & O_{n-\rho,m-\rho} \end{array}\right] , \tag{1.49}$$

where $h_1(\lambda), \dots, h_\rho(\lambda)$ are scalar monic polynomials given by the relations

$$h_1(\lambda) = a_1(\lambda), \quad h_1(\lambda)h_2(\lambda) = a_2(\lambda), \ \dots \ , h_1(\lambda)h_2(\lambda)\cdots h_\rho(\lambda) = a_\rho(\lambda) . \tag{1.50}$$

2. The polynomials $a_i(\lambda)$, $(i = 1, \dots, \rho)$ configured by (1.42) are called the *invariant polynomials* of the matrix $A(\lambda)$. It was shown that the coincidence of the sets of their invariant polynomials is not only a necessary but a sufficient condition for two polynomial matrices $A_1(\lambda)$ and $A_2(\lambda)$ to be equivalent, [51].

3. The monic greatest common divisor of all minors of i-th order for the matrix $A(\lambda)$ is named its i-th *determinantal divisor*. If $\operatorname{rank} A(\lambda) = \rho$ is true, then there exist ρ determinant divisors $D_1(\lambda)$, $D_2(\lambda)$, ... , $D_\rho(\lambda)$.

$D_\rho(\lambda)$ is named the *greatest determinantal divisor*. It can be shown that the set of determinantal divisors is invariant against equivalence transformations on the matrix $A(\lambda)$.

4. The invariant polynomials $a_i(\lambda)$ are connected with the polynomials $D_i(\lambda)$ by the relation

$$a_i(\lambda) = \frac{D_i(\lambda)}{D_{i-1}(\lambda)}\,, \quad D_0(\lambda) = 1\,, \quad (i = 1, \ldots, \rho)\,. \tag{1.51}$$

Therefore, it follows from (1.51)

$$D_1(\lambda) = a_1(\lambda),\ D_2(\lambda) = a_1(\lambda)a_2(\lambda), \quad \ldots \quad , D_\rho(\lambda) = a_1(\lambda)a_2(\lambda)\cdots a_\rho(\lambda)\,. \tag{1.52}$$

If Representations (1.49), (1.50) are used, then these relations can be written in the form

$$\begin{aligned} &D_1(\lambda) = h_1(\lambda),\ D_2(\lambda) = h_1^2(\lambda)h_2(\lambda), \quad \ldots \\ &D_\rho(\lambda) = h_1^\rho(\lambda)h_2^{\rho-1}(\lambda)\cdots h_\rho(\lambda)\,. \end{aligned}$$

5. Suppose the greatest common determinantal divisor $D_\rho(\lambda)$ be given by the linear factors

$$D_\rho(\lambda) = (\lambda - \lambda_1)^{\nu_{1\rho}} \cdots (\lambda - \lambda_q)^{\nu_{q\rho}}\,, \tag{1.53}$$

where all numbers λ_i are different. We take from (1.51) that every invariant polynomial $a_i(\lambda)$ permits a factorisation of the form

$$a_i(\lambda) = (\lambda - \lambda_1)^{\mu_{1i}} \cdots (\lambda - \lambda_q)^{\mu_{qi}}\,, \quad (i = 1, \ldots, \rho) \tag{1.54}$$

with

$$0 \le \mu_{pi} \le \mu_{p,i+1} \le \nu_{p\rho}\,, \quad (p = 1, \ldots, q)\,.$$

The factors different from one in the expression (1.54) are called *elementary divisors* of the polynomial matrix $A(\lambda)$ in the field $\mathbb{C}$. In general, every root λ_i is configured to several elementary divisors. It follows from the above said, that the set of invariant polynomials uniquely determines the set of elementary divisors. The reverse is also true if the rank of the matrix $A(\lambda)$ is known.

Example 1.22. Assume the rank of the matrix $A(\lambda)$ to be equal to four, and the whole of its elementary divisors to be

$$(\lambda - 2)^2,\ (\lambda - 2)^2,\ \lambda - 2,\ \lambda - 3,\ \lambda - 3,\ \lambda - 4.$$

Then as the set of invariant polynomials, we obtain

$$a_4(\lambda) = (\lambda-2)^2(\lambda-3)(\lambda-4),\ a_3(\lambda) = (\lambda-2)^2(\lambda-3),\ a_2(\lambda) = \lambda-2,\ a_1(\lambda) = 1.$$

Using the set of invariant polynomials, we are able to specify immediately the Smith-canonical form of the matrix $A(\lambda)$. In the present case we get

$$S_A(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda-2 & 0 & 0 \\ 0 & 0 & (\lambda-2)^2(\lambda-3) & 0 \\ 0 & 0 & 0 & (\lambda-2)^2(\lambda-3)(\lambda-4) \end{bmatrix}.$$

□

6. For diagonal and block-diagonal matrices the system of elementary divisor can be constructed by the elementary divisors of its elements.

Lemma 1.23 ([51]). *The system of elementary divisors of any diagonal matrix is the unification of the elementary divisors of its elements.* ■

Example 1.24. Let the diagonal matrix

$$A(\lambda) = \begin{bmatrix} \lambda^2 & 0 & 0 & 0 \\ 0 & \lambda(\lambda-1) & 0 & 0 \\ 0 & 0 & (\lambda-1)^2 & 0 \\ 0 & 0 & 0 & \lambda(\lambda-1) \end{bmatrix}$$

be given. By decomposition of all diagonal elements into factors (1.54), we obtain the totality of elementary divisors

$$\lambda^2,\ \lambda,\ \lambda-1,\ (\lambda-1)^2,\ \lambda,\ \lambda-1$$

and, finally, we find the Smith-canonical form

$$S_A(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda(\lambda-1) & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1) & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)^2 \end{bmatrix}.$$

□

Lemma 1.25 ([51]). *The system of elementary divisors of the block-diagonal matrix*

$$A_d(\lambda) = \begin{bmatrix} A_1(\lambda) & O & \ldots & O \\ O & A_2(\lambda) & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & A_n(\lambda) \end{bmatrix} = \operatorname{diag}\{A_1(\lambda), \ldots, A_n(\lambda)\},$$

where the $A_i(\lambda)$, $(i = 1, \ldots, n)$ *are rectangular matrices of any dimension, is built by the unification of the elementary divisors of their block elements.* ■

Example 1.26. We choose $n = 2$ and

$$A_1(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad A_2(\lambda) = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda - a \end{bmatrix}.$$

We realise immediately that the matrix $A_1(\lambda)$ possesses the one and only elementary divisor λ^3. The matrix $A_2(\lambda)$ for $a = 0$ has the two equal elementary divisors λ and λ. In case of $a \neq 0$, we find for $A_2(\lambda)$ the two different elementary divisors λ and $\lambda - a$. That's why for $a \neq 0$ the totality of elementary divisors of the matrix $A_d(\lambda) = \operatorname{diag}\{A_1(\lambda),\, A_2(\lambda)\}$ consists of $\lambda^3,\ \lambda,\ \lambda - a$, and the Smith-canonical form comes out as

$$S_{A_d}(\lambda) = \operatorname{diag}\{1,\, 1,\, 1,\, 1,\, \lambda,\, \lambda^3(\lambda - a)\}\,.$$

However, in case of $a = 0$, we find

$$S_{A_d}(\lambda) = \operatorname{diag}\{1,\, 1,\, 1,\, \lambda,\, \lambda,\, \lambda^3\}\,.$$

□

Remark 1.27. The above example illustrates the fact that the dependence of the Smith-canonical form (or the totality of its elementary divisors) on the coefficients of the polynomial matrix is numerically unstable.

1.10 Latent Equations and Latent Numbers

1. Let the non-singular matrix $A(\lambda) \in \mathbb{F}_{nn}[\lambda]$ be given. The polynomial

$$d_A(\lambda) = \det A(\lambda)$$

is said to be the *characteristic polynomial* of the matrix $A(\lambda)$, and the equation

$$d_A(\lambda) = 0 \tag{1.55}$$

is its *characteristic equation*. The roots of the characteristic equation are called the *eigenvalues* of the matrix $A(\lambda)$. For $A(\lambda) \in \mathbb{F}_{nn}[\lambda]$, the characteristic polynomial $d_A(\lambda)$ is equivalent to the greatest determinantal divisor $D_n(\lambda)$. Therefore, the characteristic equation (1.55) is equivalent to

$$D_\rho(\lambda) = 0\,, \tag{1.56}$$

where $\rho = n$ is the normal rank of the matrix $A(\lambda)$.

2. Let us consider an arbitrary matrix $A(\lambda) \in \mathbb{F}_{nm}[\lambda]$ having full rank $\rho = \gamma_A$. For it, Equation (1.56) always has a sense, and for $n \neq m$ this will be called its *latent equation*. The roots of Equation (1.56) are named *latent roots (numbers)* of the matrix $A(\lambda)$. Obviously, the latent numbers are equal to the numbers λ_i, that are configured by the factorisation (1.53). The latent roots of square matrices coincide with its eigenvalues.

Owing to (1.52), the latent equation can be written in the form

$$a_1(\lambda)a_2(\lambda)\cdots a_\rho(\lambda) = 0\,.$$

Hence it follows that every latent number is the root of at least one invariant polynomial.

3. In the following we investigate the question on the important relation between the rank of the polynomial matrix $A(\lambda)$ and the rank of the number matrix $A(\tilde{\lambda})$ that is generated from $A(\lambda)$ by substituting $\lambda = \tilde{\lambda}$, where $\tilde{\lambda}$ is a given complex number.

Theorem 1.28. *Suppose that the matrix $A(\lambda)$ possesses the rank $\rho = \gamma_A$. Then, if $\tilde{\lambda}$ does not coincide with one of the latent roots, the relation*

$$\operatorname{rank} A(\tilde{\lambda}) = \rho$$

is true. However, if $\tilde{\lambda} = \lambda_i$ takes place with a certain latent number λ_i, then we have

$$\operatorname{rank} A(\tilde{\lambda}) = \rho - d_i\,, \tag{1.57}$$

where d_i is the number of different elementary divisors, that is connected with the latent number λ_i.

Proof. If $\lambda = \tilde{\lambda}$ is not a latent number, then it follows from (1.41), (1.42)

$$\operatorname{rank} S_A(\tilde{\lambda}) = \rho\,.$$

Due to (1.40), we obtain

$$A(\tilde{\lambda}) = p(\tilde{\lambda})S_A(\lambda)q(\tilde{\lambda})\,,$$

where we read $\operatorname{rank} A(\tilde{\lambda}) = \operatorname{rank} S_A(\tilde{\lambda}) = \rho$, because the multiplication with the non-singular matrices $p(\tilde{\lambda})$, $q(\tilde{\lambda})$ does not change the rank. Now, let $\tilde{\lambda} = \lambda_i$, where λ_i is a latent number. Then there exists a number $d_i \geq 1$, such that

$$a_\rho(\lambda_i) = 0, \ldots,\ a_{\rho-d_i+1}(\lambda_i) = 0,\ a_{\rho-d_i}(\lambda_i) \neq 0, \ldots,\ a_1(\lambda_i) \neq 0\,.$$

Obviously, d_i is equal to the number of different elementary divisors of the latent root λ_i. Hence it follows with the help of (1.40)–(1.42)

$$\operatorname{rank} A(\lambda_i) = \operatorname{rank} S_A(\lambda_i) = \rho - d_i\,,$$

which is equivalent to (1.57). ∎

Corollary 1.29. *The equation for the defect*

$$\operatorname{def} A(\lambda_i) = d_i\,, \quad (i = 1, \ldots, q)$$

is true. ∎

Corollary 1.30. *It follows from Theorem 1.28, that the latent numbers λ_i of a non-degenerated matrix $A(\lambda)$ are exactly those numbers λ_i, for which*

$$\operatorname{rank} A(\lambda_i) < \gamma_A$$

becomes valid. ■

4. For a non-degenerated matrix $A(\lambda)$, let the monic greatest common divisor of the minors of γ_A-th order be equal to 1. In that case, the latent equation (1.56) has no roots, thus the matrix $A(\lambda)$ also does not possess latent roots. Such polynomial matrices are said to be *alatent*. All invariant polynomials of an alatent matrix are equal to 1.

Alatent square matrices turn out to be unimodular. For an alatent matrix $A(\lambda)$, the number matrix $A(\tilde{\lambda})$ for all $\tilde{\lambda}$ possesses its maximal rank.

Theorem 1.31. *The non-degenerated $n \times m$ matrix $A(\lambda)$ with $n < m$ proves to be alatent, if and only if there exists a unimodular matrix $\psi(\lambda)$, that meets*

$$A(\lambda) = \begin{bmatrix} I_n & O_{n,m-n} \end{bmatrix} \psi(\lambda) .$$

Proof. Under the made suppositions, due to Theorem 1.4, the Hermitian form $\tilde{A}_r(\lambda)$ has the shape

$$\tilde{A}_r(\lambda) = \begin{bmatrix} I_n & O_{n,m-n} \end{bmatrix} ,$$

and the claimed relation emerges from (1.19) for $\psi(\lambda) = q^{-1}(\lambda)$. ■

Analogously, we conclude from (1.17) that for $n > m$ the vertical $n \times m$ matrix $A(\lambda)$ is alatent if and only if

$$A(\lambda) = \varphi(\lambda) \begin{bmatrix} I_m \\ O_{n-m,m} \end{bmatrix}$$

becomes true with a certain unimodular matrix $\varphi(\lambda)$.

5. A non-degenerated matrix $A(\lambda) \in \mathbb{F}_{nm}[\lambda]$ is said to be *latent*, if it has latent roots. Due to (1.40)–(1.42) it is clear that for $n < m$ a latent matrix $A(\lambda)$ allows the representation

$$A(\lambda) = a(\lambda) b(\lambda) , \tag{1.58}$$

where

$$a(\lambda) = p(\lambda) \operatorname{diag}\{a_1(\lambda), \dots, a_n(\lambda)\} ,$$

$$b(\lambda) = \begin{bmatrix} I_n & O_{n,m-n} \end{bmatrix} q(\lambda)$$

and $p(\lambda)$, $q(\lambda)$ are unimodular matrices.

Obviously, $\det a(\lambda) \approx a_1(\lambda) \cdots a_n(\lambda)$ is valid. The matrix $b(\lambda)$ proves to be alatent, *i.e.* its rank is equal to $n = \rho$ for all $\lambda = \tilde{\lambda}$. A corresponding representation for $n > m$ is also possible.

6.

Theorem 1.32. *Suppose the $n \times m$ matrix $A(\lambda)$ to be alatent. Then every submatrix generated from any of its rows is also alatent.*

Proof. Take a positive integer $p < n$ and present the matrix $A(\lambda)$ in the form

$$A(\lambda) = \left[\begin{array}{ccc} a_{11}(\lambda) & \dots & a_{1m}(\lambda) \\ \vdots & \dots & \vdots \\ a_{p1}(\lambda) & \dots & a_{pm}(\lambda) \\ \hline a_{p+1,1}(\lambda) & \dots & a_{p+1,m}(\lambda) \\ \vdots & \dots & \vdots \\ a_{n1}(\lambda) & \dots & a_{nm}(\lambda) \end{array}\right] = \left[\begin{array}{c} A_p(\lambda) \\ \hline A_1(\lambda) \end{array}\right]. \tag{1.59}$$

It is indirectly shown that the submatrix $A_p(\lambda)$ over the line turns out to be alatent. Suppose the contrary. Then owing to (1.58), we get

$$A_p(\lambda) = a_p(\lambda) b_p(\lambda)\,,$$

where the matrix $a_p(\lambda)$ is latent, and $\operatorname{ord} a_p(\lambda) > 0$. Applying this result from (1.59),

$$A(\lambda) = \begin{bmatrix} a_p(\lambda) & O_{p,n-p} \\ O_{n-p,p} & I_{n-p} \end{bmatrix} \begin{bmatrix} b_p(\lambda) \\ A_1(\lambda) \end{bmatrix}$$

is acquired. Let $\tilde{\lambda}$ be an eigenvalue of the matrix $a_p(\lambda)$, so

$$A(\tilde{\lambda}) = \begin{bmatrix} a_p(\tilde{\lambda}) & O_{p,n-p} \\ O_{n-p,p} & I_{n-p} \end{bmatrix} \begin{bmatrix} b_p(\tilde{\lambda}) \\ A_1(\tilde{\lambda}) \end{bmatrix}$$

is valid.Because the rank of the first factors on the right side is smaller than n, this implies $\operatorname{rank} A(\tilde{\lambda}) < n$, which is in contradiction to the supposed alatency of $A(\lambda)$. ∎

Remark 1.33. In the same way, it is shown that any submatrix of an alatent matrix $A(\lambda)$ built from any of its columns also becomes alatent.

Corollary 1.34. *Every submatrix built from any rows or columns of a unimodular matrix is alatent.* ∎

1.11 Simple Matrices

1. A non-degenerated latent $n \times m$ matrix $A(\lambda)$ of full rank $\rho = \gamma_A$ is called *simple*, if

$$D_\rho(\lambda) = a_\rho(\lambda), \quad D_{\rho-1}(\lambda) = D_{\rho-2}(\lambda) = \ldots = D_1(\lambda) = 1\,.$$

In dependence on the dimension for a simple matrix $A(\lambda)$ from (1.40)–(1.42), we derive the representations

$$\boldsymbol{\gamma_A = n = m:}\ A(\lambda) = \varphi(\lambda)\,\mathrm{diag}\{1,\dots,1,a_n(\lambda)\}\psi(\lambda)\,,$$

$$\boldsymbol{\gamma_A = n < m:}\ A(\lambda) = \varphi(\lambda)\begin{bmatrix} 1 \dots 0 & 0 & 0 \dots 0 \\ \vdots \ddots \vdots & \vdots & \vdots \dots \vdots \\ 0 \dots 1 & 0 & 0 \dots 0 \\ 0 \dots 0 & a_n(\lambda) & 0 \dots 0 \end{bmatrix}\psi(\lambda)\,,$$

$$\boldsymbol{\gamma_A = m < n:}\ A(\lambda) = \varphi(\lambda)\begin{bmatrix} 1 \dots 0 & 0 \\ \vdots \ddots \vdots & \vdots \\ 0 \dots 1 & 0 \\ 0 \dots 0 & a_m(\lambda) \\ 0 \dots 0 & 0 \\ \vdots \dots \vdots & \vdots \\ 0 \dots 0 & 0 \end{bmatrix}\psi(\lambda)\,,$$

where $\varphi(\lambda)$ and $\psi(\lambda)$ are unimodular matrices.

2. From the last relations, we directly deduce the following statements:

a) For a non-degenerated matrix $A(\lambda)$ to be simple, it is necessary and sufficient that every latent root λ_i is configured to only one elementary divisor.
b) Let the non-degenerated $n \times m$ matrix $A(\lambda)$ of rank γ_A have the latent roots $\lambda_1, \dots, \lambda_q$. Then for the simplicity of $A(\lambda)$, the relation

$$\mathrm{rank}\, A(\lambda_i) = \gamma_A - 1, \quad (i = 1, \dots, q)$$

or, equivalently the condition

$$\mathrm{def}\, A(\lambda_i) = 1, \quad (i = 1, \dots, q) \tag{1.60}$$

is necessary and sufficient.
c) Another criterion for the membership of a matrix $A(\lambda)$ to the class of simple matrices yields the following theorem.

Theorem 1.35. *A necessary and sufficient condition for the simplicity of the $n \times n$ matrix $A(\lambda)$ is, that there exists a $n \times 1$ column $B(\lambda)$, such that the matrix $L(\lambda) = \begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix}$ becomes alatent.*

Proof. Sufficiency: Let the matrix $\begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix}$ be alatent and λ_i, $(i = 1, \dots, q)$ are the eigenvalues of $A(\lambda)$. Hence it follows

$$\mathrm{rank}\begin{bmatrix} A(\lambda_i) & B(\lambda_i) \end{bmatrix} = n, \quad (i = 1, \dots, q)\,.$$

Hereby, we deduce from Theorem 1.28, that we need Condition (1.60) to be satisfied if the last conditions should be fulfilled, *i.e.* the matrix $A(\lambda)$ has to be simple.

Necessity: It is shown that for a simple matrix $A(\lambda)$, there exists a column $B(\lambda)$, such that the matrix $\begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix}$ becomes alatent. Let us have $\det A(\lambda) = d(\lambda)$ and $\Delta(\lambda) \approx d(\lambda)$ as the equivalent monic polynomial. Then the matrix $A(\lambda)$ can be written in the form

$$A(\lambda) = \varphi(\lambda)\,\mathrm{diag}\{1, \dots, 1, \Delta(\lambda)\}\psi(\lambda)\,,$$

where $\varphi(\lambda)$, $\psi(\lambda)$ are unimodular $n \times n$ matrices. The matrix $Q(\lambda)$ of the shape

$$Q(\lambda) = \begin{bmatrix} I_{n-1} & O_{n-1,1} & O_{n-1,1} \\ O_{1,n-1} & \Delta(\lambda) & 1 \end{bmatrix}$$

is obviously alatent, because it has a minor of n-th order that is equal to one. The matrix $\tilde{\psi}(\lambda)$ with

$$\tilde{\psi}(\lambda) = \begin{bmatrix} \psi(\lambda) & O_{n1} \\ O_{1n} & 1 \end{bmatrix} = \mathrm{diag}\{\psi(\lambda), 1\}$$

is unimodular. Applying the last two equations, we get

$$\varphi(\lambda)Q(\lambda)\tilde{\psi}(\lambda) = \begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix} = L(\lambda)$$

with

$$B(\lambda) = \varphi(\lambda) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} . \tag{1.61}$$

The matrix $L(\lambda)$ is alatent per construction. ∎

Remark 1.36. If the matrix $\varphi(\lambda)$ is written in the form

$$\varphi(\lambda) = \begin{bmatrix} \varphi_1(\lambda) & \dots & \varphi_n(\lambda) \end{bmatrix},$$

where $\varphi_1(\lambda), \dots, \varphi_n(\lambda)$ are the corresponding columns, then from (1.61), we gain

$$B(\lambda) = \varphi_n(\lambda)\,.$$

3. Square simple matrices possess the property of structural stability, which will be explained by the next theorem.

Theorem 1.37. *Let the matrices* $A(\lambda) \in \mathbb{F}_{nn}(\lambda)$, $B(\lambda) \in \mathbb{F}_{nn}[\lambda]$ *be given, where the matrix* $A(\lambda)$ *is simple, but the matrix* $B(\lambda)$ *is of any structure. Furthermore, let us have* $\det A(\lambda) = d(\lambda)$ *and*

$$\det[A(\lambda) + \epsilon B(\lambda)] = d(\lambda) + \epsilon d_1(\lambda, \epsilon)\,, \tag{1.62}$$

where $d_1(\lambda, \epsilon)$ is a polynomial, satisfying the condition

$$\deg d_1(\lambda, \epsilon) < \deg d(\lambda)\,. \tag{1.63}$$

Then there exists a positive number ϵ_0, such that for $|\epsilon| < \epsilon_0$ all matrices $A(\lambda) + \epsilon B(\lambda)$ are simple.

Proof. The proof splits into several stages.

Lemma 1.38. *Let $\|\cdot\|$ be a certain norm for finite-dimensional number matrices. Then for any matrix $B = [\,b_{ik}\,] \in \mathbb{F}_{nn}$ the estimation*

$$\max_{1 \le i,k \le n} |b_{ik}| \le \beta \|B\| \tag{1.64}$$

is true, where $\beta > 0$ is a constant, independent of B.

Proof. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be any two norms in the space $\mathbb{C}_{nn}$. Due to the finite dimension of $\mathbb{C}_{nn}$, any two norms are equivalent, that means, for an arbitrary matrix B, we have

$$\alpha_1 \|B\|_1 \le \|B\|_2 \le \alpha_2 \|B\|_1\,,$$

where α_1, α_2 are positive constants not depending on the choice of B.

Take

$$\|B\|_1 = \max_{1 \le i \le n} \sum_{k=1}^{n} |b_{ik}|\,,$$

then under the assumption $\|\cdot\|_2 = \|\cdot\|$, we win

$$|b_{ik}| \le \|B\|_1 \le \alpha_1^{-1} \|B\|\,,$$

which is adequate to (1.64) with $\beta = \alpha_1^{-1}$. ∎

Lemma 1.39. *Let the matrix $A \in \mathbb{F}_{nn}$ be non-singular and $\|\cdot\|$ be a certain norm in $\mathbb{F}_{nn}$. Then, there exists a positive constant α_0, such that for $\|B\| < \alpha_0$, all matrices $A + B$ become non-singular.*

Proof. Assume $|b_{ik}| \le \beta \|B\|$, where $\beta > 0$ is the constant configured in (1.64). Then we expand

$$\det(A + B) = \det A + \varphi(A, B)\,,$$

where $\varphi(A, B)$ is a scalar function of the elements in A and B. For it, an estimation

$$|\varphi(A, B)| < \mu_1 \beta \|B\| + \mu_2 \beta^2 \|B\|^2 + \ldots + \mu_n \beta^n \|B\|^n$$

is true, where μ_i, $(i = 1, \ldots, n)$ are constants, that do not depend on B. Hence there exists a number $\alpha_0 > 0$, such that $\|B\| < \alpha_0$ always implies

$$|\varphi(A, B)| < |\det A|\,.$$

That's why for $\|B\| < \alpha_0$, the desired relation $\det(A + B) \ne 0$ holds. ∎

Lemma 1.40. *For the matrix $A \in \mathbb{F}_{nn}$, we assume* $\operatorname{rank} A = \rho$, *and let* $\|\cdot\|$ *be a certain norm in $\mathbb{F}_{nn}$. Then there exists a positive constant α_0, such that for $B \in \mathbb{F}_{nn}$ with $\|B\| < \alpha_0$ always*

$$\operatorname{rank}(A+B) \geq \rho\,.$$

Proof. Let A_ρ be a non-zero minor of order ρ of A. Lemma 1.39 delivers the existence of a number $\alpha_0 > 0$, such that for $\|B\| < \alpha_0$, the minor of the matrix $A+B$ corresponding to A_ρ is different from zero. However, this means that the rank will not reduce after addition of B, but that was claimed by the lemma. ∎

Proof of Theorem 1.37 Let

$$d(\lambda) = d_0\lambda^k + \ldots + d_k = 0, \quad d_0 \neq 0$$

be the characteristic polynomial of the matrix $A(\lambda)$. Then we obtain from (1.62) and (1.63)

$$\det[A(\lambda) + \epsilon B(\lambda)] \stackrel{\triangle}{=} d(\lambda, \epsilon) = d_0\lambda^k + d_1(\epsilon)\lambda^{k-1} + \ldots + d_k(\epsilon)\,,$$

where

$$d_i(\epsilon) = d_i + d_{i1}\epsilon + d_{i2}\epsilon^2 + \ldots \quad , \; (i = 1, \ldots, k)$$

are polynomials in the variable ϵ with $d_i(0) = d_i$. Let $\tilde{\lambda}$ be a root of the equation

$$d(\lambda, 0) = d(\lambda) = 0$$

with multiplicity ν, *i.e.* an eigenvalue of the matrix $A(\lambda)$ with multiplicity ν. Since the matrix $A(\lambda)$ is simple, we obtain

$$\operatorname{rank} A(\tilde{\lambda}) = n - 1\,.$$

Hereby, due to Lemma 1.40 it follows the existence of a constant $\tilde{\alpha}$, such that for every matrix $G \in \mathbb{C}_{nn}$ with $\|G\| < \tilde{\alpha}$ the relation

$$\operatorname{rank}[A(\tilde{\lambda}) + G] \geq n - 1 \tag{1.65}$$

is fulfilled. Now consider the equation

$$d(\lambda, \epsilon) = 0\,.$$

As known from [188], for $|\epsilon| < \delta$, where $\delta > 0$ is sufficiently small, there exist ν continuous functions $\tilde{\lambda}_i(\epsilon)$, $(i = 1, \ldots, \nu)$, such that

$$d(\tilde{\lambda}_i(\epsilon), \epsilon) = \det[A(\tilde{\lambda}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i(\epsilon))] = 0\,, \tag{1.66}$$

where some of the functions $\tilde{\lambda}_i(\epsilon)$ may coincide. Thereby, the limits

$$\lim_{\epsilon \to 0} \tilde{\lambda}_i(\epsilon) = \tilde{\lambda}_i , \quad (i = 1, \dots, \nu)$$

exist, and we can write

$$\tilde{\lambda}_i(\epsilon) = \tilde{\lambda}_i + \tilde{\psi}_i(\epsilon) ,$$

where $\tilde{\psi}_i(\epsilon)$ are continuous functions with $\tilde{\psi}(0) = 0$. Consequently, we get

$$\begin{aligned} A(\tilde{\lambda}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i(\epsilon)) &= A(\tilde{\lambda}_i + \tilde{\psi}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i + \tilde{\psi}_i(\epsilon)) \\ &= A(\tilde{\lambda}_i) + G_i(\epsilon) \end{aligned}$$

with

$$G_i(\epsilon) = \epsilon B(\tilde{\lambda}_i) + \tilde{L}_i(\epsilon)$$

and the matrices $\tilde{L}_i(\epsilon)$ for $|\epsilon| < \delta$ depend continuously on ϵ, and $\tilde{L}_i(0) = O_{nn}$ holds. Next choose a constant $\tilde{\epsilon} > 0$ with the property that for $|\epsilon| < \tilde{\epsilon}$ and all $i = 1, \dots, \nu$, the relation

$$\|G_i(\epsilon)\| = \|\epsilon B(\tilde{\lambda}_i) + \tilde{L}_i(\epsilon)\| < \tilde{\alpha}$$

is true. Therefore, we receive for $|\epsilon| < \tilde{\epsilon}$ from (1.65)

$$\operatorname{rank}[A(\tilde{\lambda}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i(\epsilon))] \geq n - 1 .$$

On the other side, it follows from (1.66), that for $|\epsilon| < \delta$, we have

$$\operatorname{rank}[A(\tilde{\lambda}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i(\epsilon))] \leq n - 1 .$$

Comparing the last two inequalities, we find for $|\epsilon| < \min\{\tilde{\epsilon}, \delta\}$

$$\operatorname{rank}[A(\tilde{\lambda}_i(\epsilon)) + \epsilon B(\tilde{\lambda}_i(\epsilon))] = n - 1 .$$

The above considerations can be made for all eigenvalues of the matrix $A(\lambda)$, therefore, Theorem 1.37 is proved by (1.60). ■

1.12 Pairs of Polynomial Matrices

1. Let us have $a(\lambda) \in \mathbb{F}_{nn}[\lambda]$, $b(\lambda) \in \mathbb{F}_{nm}[\lambda]$. The entirety of both matrices is called a *horizontal pair*, and it is designated by $(a(\lambda), b(\lambda))$. On the other side, if we have $a(\lambda) \in \mathbb{F}_{mm}[\lambda]$ and $c(\lambda) \in \mathbb{F}_{nm}[\lambda]$, then we speak about a *vertical pair* and we write $[a(\lambda), c(\lambda)]$. The pairs $(a(\lambda), b(\lambda))$ and $[a(\lambda), c(\lambda)]$ may be configured to the rectangular matrices

$$R_h(\lambda) = \begin{bmatrix} a(\lambda) & b(\lambda) \end{bmatrix} , \quad R_v(\lambda) = \begin{bmatrix} a(\lambda) \\ c(\lambda) \end{bmatrix} , \tag{1.67}$$

where the first one is horizontal, and the second one is vertical. Due to

$$R'_v(\lambda) = \left[\, a'(\lambda) \; c'(\lambda) \,\right],$$

the properties of vertical pairs can immediately deduced from the properties of horizontal pairs. Therefore, we will consider now only horizontal pairs. The pairs $(a(\lambda), b(\lambda))$, $[a(\lambda), c(\lambda)]$ are called *non-degenerated* if the matrices (1.67) are non-degenerated. If not supposed explicitly otherwise, we will always consider non-degenerated pairs.

2. Let for the pair $(a(\lambda), b(\lambda))$ exist a polynomial matrix $g(\lambda)$, such that

$$a(\lambda) = g(\lambda)a_1(\lambda)\,, \quad b(\lambda) = g(\lambda)b_1(\lambda) \tag{1.68}$$

with polynomial matrices $a_1(\lambda)$, $b_1(\lambda)$. Then the matrix $g(\lambda)$ is called a *common left divisor* of the pair $(a(\lambda), b(\lambda))$. The common left divisor $g(\lambda)$ is named as a *greatest common left divisor* (GCLD) of the pair $(a(\lambda), b(\lambda))$, if for any left common divisor $g_1(\lambda)$

$$g(\lambda) = g_1(\lambda)\alpha(\lambda)$$

with a polynomial matrix $\alpha(\lambda)$ is true. As known any two GCLD are right-equivalent [69].

3. If the pair $(a(\lambda), b(\lambda))$ is non-degenerated, then from Theorem 1.4, it follows the existence of a unimodular matrix

$$r(\lambda) = \begin{array}{cc} \begin{array}{cc} \quad n & \qquad m \end{array} & \\ \begin{bmatrix} r_{11}(\lambda) & r_{12}(\lambda) \\ r_{21}(\lambda) & r_{22}(\lambda) \end{bmatrix} & \begin{array}{c} n \\ m \end{array} \end{array} \tag{1.69}$$

for which

$$\left[\, a(\lambda) \; b(\lambda) \,\right] \begin{bmatrix} r_{11}(\lambda) \; r_{12}(\lambda) \\ r_{21}(\lambda) \; r_{22}(\lambda) \end{bmatrix} = \begin{array}{cc} \begin{array}{cc} \;\; n & \quad m \end{array} & \\ \left[\, N(\lambda) \; O \,\right] & n \end{array} \tag{1.70}$$

holds. As known [69], the matrix $N(\lambda)$ is a GCLD of the pair $(a(\lambda), b(\lambda))$.

4. The pair $(a(\lambda), b(\lambda))$ is called *irreducible*, if the matrix $R_h(\lambda)$ in (1.67) is alatent. From the above considerations, it follows that the pair $(a(\lambda), b(\lambda))$ is irreducible, if and only if there exists a unimodular matrix $r(\lambda)$ according to (1.69) with

$$\left[\, a(\lambda) \; b(\lambda) \,\right] r(\lambda) = \left[\, I_n \; O_{nm} \,\right] .$$

5. Let

$$s(\lambda) = r^{-1}(\lambda) = \begin{bmatrix} s_{11}(\lambda) \; s_{12}(\lambda) \\ s_{21}(\lambda) \; s_{22}(\lambda) \end{bmatrix}$$

be a unimodular polynomial matrix. Then we get from (1.70)

$$\begin{bmatrix} a(\lambda) & b(\lambda) \end{bmatrix} = \begin{bmatrix} N(\lambda) & O_{nm} \end{bmatrix} \begin{bmatrix} s_{11}(\lambda) & s_{12}(\lambda) \\ s_{21}(\lambda) & s_{22}(\lambda) \end{bmatrix}.$$

Hence it follows immediately

$$a(\lambda) = N(\lambda)s_{11}(\lambda), \quad b(\lambda) = N(\lambda)s_{12}(\lambda),$$

that can be written in the form

$$\begin{bmatrix} a(\lambda) & b(\lambda) \end{bmatrix} = N(\lambda) \begin{bmatrix} s_{11}(\lambda) & s_{12}(\lambda) \end{bmatrix}.$$

Due to Corollary 1.34, the pair $(s_{11}(\lambda), s_{12}(\lambda))$ is irreducible. Therefore, the next statement is true:

If Relation (1.68) is true, and $g(\lambda)$ is a GCLD of the pair $(a(\lambda), b(\lambda))$, then the pair $(a_1(\lambda), b_1(\lambda))$ is irreducible.

The reverse statement is also true:

If Relation (1.68) is valid, and the pair $(a_1(\lambda), b_1(\lambda))$ is irreducible, then the matrix $g(\lambda)$ is a GCLD of the pair $(a(\lambda), b(\lambda))$.

6. A necessary and sufficient condition for the irreducibility of the pair $(a(\lambda), b(\lambda))$ with the $n \times n$ polynomial matrix $a(\lambda)$ and the $n \times m$ polynomial matrix $b(\lambda)$ is the existence of an $n \times n$ polynomial matrix $X(\lambda)$ and an $m \times n$ polynomial matrix $Y(\lambda)$, such that the relation

$$a(\lambda)X(\lambda) + b(\lambda)Y(\lambda) = I_n \tag{1.71}$$

becomes true [69].

7. All what is said up to now, can be transferred practically without change to vertical pairs $[a(\lambda), c(\lambda)]$. In this case, instead of the concepts common left divisor and GCLD we introduce the concepts *common right divisor* and *greatest common right divisor* (GCRD) . Hereby, if

$$p(\lambda)\begin{bmatrix} a(\lambda) \\ c(\lambda) \end{bmatrix} = \begin{array}{cc} m & \\ \begin{bmatrix} L(\lambda) \\ O_{nm} \end{bmatrix} & \begin{array}{c} m \\ n \end{array} \end{array}$$

is valid with a unimodular matrix $p(\lambda)$, then $L(\lambda)$ is a GCRD of the corresponding pair $[a(\lambda), c(\lambda)]$. If $L(\lambda)$ and $L_1(\lambda)$ are two GCRD, then they are related by

$$L(\lambda) = f(\lambda)L_1(\lambda)$$

where $f(\lambda)$ is a unimodular matrix.

The vertical pair $[a(\lambda), c(\lambda)]$ is called *irreducible*, if the matrix $R_v(\lambda)$ in (1.67) is alatent. The pair $[a(\lambda), c(\lambda)]$ turns out to be irreducible, if and only if, there exists a unimodular matrix $p(\lambda)$ with

$$p(\lambda)\begin{bmatrix} a(\lambda) \\ c(\lambda) \end{bmatrix} = \begin{bmatrix} I_m \\ O_{nm} \end{bmatrix}.$$

Immediately, it is seen that the pair $[a(\lambda), c(\lambda)]$ is exactly irreducible, when there exist polynomial matrices $U(\lambda)$, $V(\lambda)$, for which

$$U(\lambda)a(\lambda) + V(\lambda)c(\lambda) = I_m\,.$$

8. The above stated irreducibility criteria will be formulated alternatively.

Theorem 1.41. *A necessary and sufficient condition for the pair $(a(\lambda), b(\lambda))$ to be irreducible, is the existence of a pair $(\alpha_l(\lambda), \beta_l(\lambda))$, such that the matrix*

$$Q_l(\lambda) = \begin{bmatrix} a(\lambda) & b(\lambda) \\ \beta_l(\lambda) & \alpha_l(\lambda) \end{bmatrix}$$

becomes unimodular.

For the pair $[a(\lambda), c(\lambda)]$ to be irreducible, it is necessary and sufficient that there exists a pair $[\alpha_r(\lambda), \beta_r(\lambda)]$, such that the matrix

$$Q_r(\lambda) = \begin{bmatrix} \alpha_r(\lambda) & c(\lambda) \\ \beta_r(\lambda) & a(\lambda) \end{bmatrix}$$

becomes unimodular. ■

9.

Lemma 1.42. *Necessary and sufficient for the irreducibility of the pair $(a(\lambda), b(\lambda))$, with the $n \times n$ and $n \times m$ polynomial matrices $a(\lambda)$ and $b(\lambda)$, is the condition*

$$\operatorname{rank} R_h(\lambda_i) = \operatorname{rank}\begin{bmatrix} a(\lambda_i) & b(\lambda_i) \end{bmatrix} = n\,, \qquad (i = 1, \ldots, q)\,, \tag{1.72}$$

where the λ_i are the different eigenvalues of the matrix $a(\lambda)$.

Proof. Sufficiency: For $\lambda = \tilde{\lambda} \neq \lambda_i$, $(i = 1, \ldots, q)$, we have $\operatorname{rank} a(\tilde{\lambda}) = n$. Therefore, together with (1.72) the relation $\operatorname{rank} R_h(\lambda) = n$ is true for all finite λ. This means, however, the pair $(a(\lambda), b(\lambda))$ is irreducible.
The *necessity* of Condition (1.72) is obvious. ■

10.

Lemma 1.43. *Let the pair $(a(\lambda), b(\lambda))$ be given with the $n \times n$ and $n \times m$ polynomial matrices $a(\lambda)$, $b(\lambda)$. Then for the pair $(a(\lambda), b(\lambda))$ to be irreducible, it is necessary that the matrix $a(\lambda)$ has not more than m invariant polynomials different from 1.*

Proof. Assume the number of invariant polynomials different from 1 of the matrix $a(\lambda)$ be $\kappa > m$. Then it follows from (1.57), that there exists an eigenvalue λ_0 of the matrix $a(\lambda)$ with $\operatorname{rank} a(\lambda_0) = n - \kappa$. Applying Inequality (1.48), we gain $\operatorname{rank}\left[\, a(\lambda_0)\; b(\lambda_0) \,\right] \leq n - \kappa + m < n$ that means, the matrix $\left[\, a(\lambda)\; b(\lambda) \,\right]$ is not alatent and, consequently, the pair $(a(\lambda), b(\lambda))$ is not irreducible. ∎

Remark 1.44. Obviously, we could formulate adequate statements as in Lemmata 1.42 and 1.43 for vertical pairs too.

1.13 Polynomial Matrices of First Degree (Pencils)

1. For $q = 1,\ n = m$ the polynomial matrix (1.10) takes the form

$$A(\lambda) = A\lambda + B \tag{1.73}$$

with constant $n \times n$ matrices A, B. This special structure is also called a *pencil.* The pencil $A(\lambda)$ is non-singular if

$$\det(A\lambda + B) \not\equiv 0\,.$$

According to the general definition, the non-singular matrix (1.73) is called *regular* for $\det A \neq 0$ and *anomalous* for $\det A = 0$. Regular pencils arise in connection with state space representations, while anomalous pencils are configured to descriptor systems [109, 34, 182]. All introduced concepts and statements that were developed for polynomial matrices of general structure are also valid for pencils (1.73). At the same time, these matrices possess a number of important additional properties that will be investigated in this section. In what follows, we only consider non-singular pencils.

2. In accordance with the general definition, the two matrices of equal dimension

$$A(\lambda) = A\lambda + B, \quad A_1(\lambda) = A_1\lambda + B_1 \tag{1.74}$$

are called left(right)-equivalent, if there exists a unimodular matrix $p(\lambda)$ ($q(\lambda)$), such that

$$A(\lambda) = p(\lambda)A_1(\lambda), \qquad (A(\lambda) = A_1(\lambda)q(\lambda))\,.$$

The matrices (1.74) are equivalent, if they satisfy an equation

$$A(\lambda) = p(\lambda)A_1(\lambda)q(\lambda)$$

with unimodular matrices $p(\lambda)$, $q(\lambda)$. As follows from the above disclosures, the matrices (1.74) are exactly left(right)-equivalent, if their Hermitian canonical forms coincide. For the equivalence of the matrices (1.74), it is necessary and sufficient that their Smith-canonical forms coincide.

3. The matrices (1.74) are named *strictly equivalent*, if there exist constant matrices P, Q with

$$A(\lambda) = PA_1(\lambda)Q\,. \tag{1.75}$$

If in (1.74) the conditions $\det A \neq 0$, $\det A_1 \neq 0$ are valid, *i.e.* the matrices are regular, then the matrices $A(\lambda)$, $B(\lambda)$ are only in that case equivalent, when they are strictly equivalent. If $\det A = 0$ or $\det A_1 = 0$, *i.e.* the matrices (1.74) are anomalous, then the conditions for equivalence and strict equivalence do not coincide.

4. In order to formulate a criterion for the strict equivalence of anomalous matrices (1.74), following [51], we consider the $n \times n$ *Jordan block*

$$J_n(a) \triangleq \begin{bmatrix} a & 1 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & \dots & 0 & 0 \\ 0 & 0 & a & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}, \tag{1.76}$$

where a is a constant.

Theorem 1.45 ([51]). *Let*

$$\det A(\lambda) = \det(A\lambda + B) \neq 0$$

be given with $\det A = 0$ *and*

$$0 < \operatorname{ord} A(\lambda) = \deg \det A(\lambda) = \eta < n\,. \tag{1.77}$$

Furthermore, let

$$(\lambda - \lambda_1)^{\eta_1}, \dots, (\lambda - \lambda_q)^{\eta_q}, \qquad \eta_1 + \dots + \eta_q = \eta \tag{1.78}$$

be the entity of elementary divisors of $A(\lambda)$ in the field $\mathbb{C}$. In what follows, the elementary divisors (1.78) will be called finite elementary divisors. *Then the matrix $A(\lambda)$ is strictly equivalent to the matrix*

$$\tilde{A}(\lambda) = \operatorname{diag}\{\lambda I_\eta + A_\eta, I_{n-\eta} + \lambda A_\nu\} \tag{1.79}$$

with

$$\begin{aligned} A_\eta &= \operatorname{diag}\{J_{\eta_1}(\lambda_1), \dots, J_{\eta_q}(\lambda_q)\}\,, \\ A_\nu &= \operatorname{diag}\{J_{p_1}(0), \dots, J_{p_\ell}(0)\}\,, \end{aligned} \tag{1.80}$$

where $p_1, \dots, p_\ell$ are positive integers with $p_1 + \dots + p_\ell = n - \eta$. The matrix $A_\nu \in \mathbb{C}_{n-\eta,n-\eta}$ is nilpotent, that means, there exists an integer κ with $A_\nu^\kappa = O_{n-\eta,n-\eta}$. ■

Remark 1.46. The above defined numbers $p_1, \dots, p_\ell$ are determined by the infinite elementary divisors of the matrix $A(\lambda)$, [51]. Thereby, the matrices (1.74) are strictly equivalent, if their finite and infinite elementary divisors coincide.

Remark 1.47. Matrix (1.79) can be represented as

$$\tilde{A}(\lambda) = U\lambda + V\,, \tag{1.81}$$

where

$$U = \operatorname{diag}\{I_\eta, A_\nu\}\,, \quad V = \operatorname{diag}\{A_\eta, I_{n-\eta}\}\,. \tag{1.82}$$

As is seen from (1.76) and (1.80)–(1.82) for $\eta < n$, we always obtain $\det U = 0$ and the matrix $\tilde{A}(\lambda)$ is generally spoken not row reduced.

5. As any non-singular matrix, also an anomalous matrix (1.73) can be brought into row reduced form by left equivalence transformations. Hereby, we obtain for matrices of first degree some further results.

Theorem 1.48. *Let Relation (1.77) be true for the non-singular anomalous matrix (1.73). Then there exists a unimodular matrix $P(\lambda)$, such that*

$$P(\lambda)(A\lambda + B) = \tilde{A}(\lambda) = \tilde{A}\lambda + \tilde{B} \tag{1.83}$$

is true with constant matrices

$$\tilde{A} = \begin{array}{c} n \\ \left[\begin{array}{c} \tilde{A}_1 \\ O_{n-\eta,n} \end{array}\right] \end{array} \begin{array}{l} \eta \\ n-\eta \end{array}\,, \quad \tilde{B} = \begin{array}{c} n \\ \left[\begin{array}{c} \tilde{B}_1 \\ \tilde{B}_2 \end{array}\right] \end{array} \begin{array}{l} \eta \\ n-\eta \end{array}\,. \tag{1.84}$$

Moreover

$$\det \begin{bmatrix} \tilde{A}_1 \\ \tilde{B}_2 \end{bmatrix} \neq 0 \tag{1.85}$$

is true together with

$$\deg P(\lambda) \le n - \eta\,. \tag{1.86}$$

Proof. We apply the row transformation algorithm of Theorem 1.9 to the matrix $A\lambda + B$. Then after a finite number of steps, we reach at a row reduced matrix $\tilde{A}(\lambda)$. Due to the fact, that the degree of the transformed matrix does not increase, we conclude $\deg \tilde{A}(\lambda) \le 1$. The case $\deg \tilde{A}(\lambda) = 0$ is excluded, otherwise the matrix $A(\lambda)$ would be unimodular in contradiction to (1.77). Therefore, only $\deg \tilde{A}(\lambda) = 1$ is possible. Moreover, we prove

$$\tilde{A}(\lambda) = \tilde{A}\lambda + \tilde{B} = \operatorname{diag}\{\lambda^{\alpha_1}, \dots, \lambda^{\alpha_n}\}\left(\tilde{A}_0 + \tilde{A}_1\lambda^{-1}\right) \tag{1.87}$$

with $\det \tilde{A}_0 \neq 0$, where each of the numbers α_i, $(i = 1, \dots, n)$ is either 0 or 1. Due to

$$\alpha_1 + \ldots + \alpha_n = \eta\,,$$

among the numbers $\alpha_1, \ldots, \alpha_n$ are exactly η ones with the value one, and the other $n - \eta$ numbers are zero. Without loss of generality, we assume the succession

$$\alpha_1 = \alpha_2 = \ldots = \alpha_\eta = 1, \quad \alpha_{\eta+1} = \alpha_{\eta+2} = \ldots = \alpha_n = 0\,.$$

Then the matrix $\tilde{A}$ in (1.83) takes the shape (1.84). Furthermore, if the matrix $\tilde{A}(\lambda)$ is represented in the form (1.87), then with respect to (1.79) and (1.84), we get

$$\tilde{A}_0 = \begin{bmatrix} \tilde{A}_1 \\ \tilde{B}_2 \end{bmatrix}.$$

Since the matrix $\tilde{A}(\lambda)$ is row reduced, Relation (1.85) arises.

It remains to show Relation (1.86). As follows from (1.36), each step decreases the degree of one of the rows of the transformed matrices at least by one. Hence each row of the matrix $A(\lambda)$ cannot be transformed more than once. Therefore, the number of transformation steps is at most $n - \eta$. Since however, in every step the transformation matrix $P(\lambda)$ is either constant or with degree one, Relation (1.86) holds. ■

Corollary 1.49. *In the row-reduced form (1.83), $n - \eta$ rows of the matrix $\tilde{A}(\lambda)$ are constant. Moreover, the rank of the matrix built from these rows is equal to $n - \eta$, i.e., these rows are linearly independent.* ■

Example 1.50. Consider the anomalous matrix

$$A(\lambda) = A\lambda + B = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \lambda + \begin{bmatrix} 2 & 1 & 3 \\ 3 & 2 & 5 \\ 3 & 2 & 6 \end{bmatrix} = \begin{bmatrix} \lambda+2 & \lambda+1 & 2\lambda+3 \\ \lambda+3 & \lambda+2 & 2\lambda+5 \\ \lambda+3 & \lambda+2 & 3\lambda+6 \end{bmatrix}$$

appering in [51], that is represented in the form

$$A(\lambda) = \operatorname{diag}\{\lambda, \lambda, \lambda\} \left(A_0 + A_1\lambda^{-1}\right)$$

with

$$A_0 = A, \quad A_1 = B\,.$$

In the first transformation step (1.30), we obtain

$$\nu_1 + \nu_2 + \nu_3 = 0$$
$$2\nu_1 + 2\nu_2 + 3\nu_3 = 0\,.$$

Now, we can choose $\nu_1 = 1$, $\nu_2 = -1$, $\nu_3 = 0$, and the matrices (1.32) and (1.35) take the form

$$P_1(\lambda) = D_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

hence

$$A_1(\lambda) = P_1(\lambda)A(\lambda) = \operatorname{diag}\{1, \lambda, \lambda\} \left(\begin{bmatrix} -1 & -1 & -2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 3 & 2 & 5 \\ 3 & 2 & 6 \end{bmatrix} \lambda^{-1} \right).$$

By appropriate manipulations, these matrices are transformed into

$$P_2(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Finally, we receive over the product

$$A_2(\lambda) = P_2(\lambda)A_1(\lambda) = P_2(\lambda)P_1(\lambda)A(\lambda)$$

the row-reduced matrix

$$A_2(\lambda) = \begin{bmatrix} -1 & -1 & -2 \\ 3 & 2 & 5 \\ \lambda + 3 & \lambda + 2 & 3\lambda + 6 \end{bmatrix}.$$

□

6. Let B be a constant $n \times n$ matrix. We assign to this matrix a matrix B_λ of degree one by

$$B_\lambda = \lambda I_n - B\,,$$

which is called the *characteristic matrix* of B. For polynomial matrices of this form, all above introduced concepts and statements for polynomial matrices of general form remain valid. Hereby, the characteristic polynomial of the matrix B_λ

$$\det B_\lambda = \det(\lambda I_n - B) \stackrel{\triangle}{=} d_B(\lambda)$$

usually is named the *characteristic polynomial* of the matrix B. In the same way, we deal with the terminology of minimal polynomials, invariant polynomials, elementary divisor *etc.* Obviously,

$$\operatorname{ord} B_\lambda = \deg \det B_\lambda = n\,.$$

As a consequence from Relation (1.75) for $A_1 = A = I_n$ we formulate:

Theorem 1.51 ([51]). *For two characteristic matrices $B_\lambda = \lambda I_n - B$ and $B_{1\lambda} = \lambda I_n - B_1$ to be equivalent, it is necessary and sufficient, that the matrices B and B_1 are similar, i.e. the relation*

$$B_1 = LBL^{-1}$$

is true with a certain non-singular constant matrix L. ■

Remark 1.52. Theorem 1.51 implies the following property. If the matrix B (the matrix $\lambda I_n - B$) has the entirety of elementary divisors

$$(\lambda - \lambda_1)^{\nu_1} \cdots (\lambda - \lambda_q)^{\nu_q}, \quad \nu_1 + \ldots + \nu_q = n,$$

then the matrix B is similar to the matrix J of the form

$$J = \operatorname{diag}\{J_{\nu_1}(\lambda_1), \ldots, J_{\nu_q}(\lambda_q)\}. \tag{1.88}$$

The matrix J is said to be the *Jordan (canonical) form* or shortly, *Jordan matrix* of the corresponding matrix B. For any $n \times n$ matrix B, the Jordan matrix is uniquely determined, except the succession of the diagonal blocks.

7. Let the horizontal pair of constant matrices (A, B) with A, $n \times n$, and B, $n \times m$ be given. The pair (A, B) is called *controllable*, if the polynomial pair $(\lambda I_n - A, B)$ is irreducible. This means, that the pair (A, B) is controllable if and only if the matrix

$$R_c(\lambda) = \begin{bmatrix} \lambda I_n - A & B \end{bmatrix}$$

is alatent. It is known, see for instance [72, 69], that the pair (A, B) is controllable, if and only if

$$\operatorname{rank} Q_c(A, B) = n,$$

where the matrix $Q_c(A, B)$ is determined by

$$Q_c(A, B) = \begin{bmatrix} B \; AB \; \ldots \; A^{n-1}B \end{bmatrix}. \tag{1.89}$$

The matrix $Q_c(A, B)$ is named *controllability matrix* of the pair (A, B).

Some statements regarding the controllability of pairs are listed now:

a) If the pair (A, B) is controllable, and the $n \times n$ matrix R is non-singular, then also the pair (A_1, B_1) with $A_1 = RAR^{-1}$, $B_1 = RB$ is controllable. Indeed, from (1.89) we obtain

$$Q_c(A_1, B_1) = \begin{bmatrix} RB \; RAB \; \ldots \; RA^{n-1}B \end{bmatrix} == RQ_c(A, B),$$

from which follows $\operatorname{rank} Q_c(A_1, B_1) = \operatorname{rank} Q_c(A, B) = n$, because R is non-singular.

b) **Theorem 1.53.** *Let the pair (A, B) with the $n \times n$ matrix A and the $n \times m$ matrix B be given, and moreover, an $n \times n$ matrix L, which is commutative with A, i.e. $AL = LA$. Then the following statements are true:*
   1. *If the pair (A, B) is not controllable, then the pair (A, LB) is also not controllable.*
   2. *If the pair (A, B) is controllable and the matrix L is non-singular, then the pair (A, LB) is controllable.*
   3. *If the matrix L is singular, then the pair (A, LB) is not controllable.*

Proof. The controllability matrix of the pair (A, LB) has the shape

$$\begin{aligned} Q_c(A, LB) &= \left[\, LB \ ALB \ \dots \ A^{n-1}LB \,\right] \\ &= L \left[\, B \ AB \ \dots \ A^{n-1}B \,\right] = LQ_c(A, B) \end{aligned} \tag{1.90}$$

where $Q_c(A, B)$ is the controllability matrix (1.89). If the pair (A, B) is not controllable, then we have $\operatorname{rank} Q_c(A, B) < n$, and therefore, $\operatorname{rank} Q_c(A, LB) < n$. Thus the 1st statement is proved. If the pair (A, B) is controllable and the matrix L is non-singular, then we have $\operatorname{rank} Q_c(A, B) = n$, $\operatorname{rank} L = n$ and from (1.90) it follows $\operatorname{rank} Q_c(A, LB) = n$. Hence the 2nd statement is shown. Finally, if the matrix L is singular, then $\operatorname{rank} L < n$ and $\operatorname{rank} Q_c(A, LB) < n$ are true, which proves 3. ∎

c) Controllable pairs are structural stable - this is stated in the next theorem.

Theorem 1.54. *Let the pair (A, B) be controllable, and (A_1, B_1) be an arbitrary pair of the same dimension. Then there exists a positive number ϵ_0, such that the pair $(A + \epsilon A_1, B + \epsilon B_1)$ is controllable for all $|\epsilon| < \epsilon_0$.*

Proof. Using (1.89) we obtain

$$Q_c(A + \epsilon A_1, B + \epsilon B_1) = Q_c(A, B) + \epsilon Q_1 + \ldots + \epsilon^n Q_n \,, \tag{1.91}$$

where the Q_i, $(i = 1, \ldots, n)$ are constant matrices, that do not depend on ϵ. Since the pair (A, B) is controllable, the matrix $Q_c(A, B)$ contains a non-zero minor of n-th order. Then due to Lemma 1.39 for sufficiently small $|\epsilon|$, the corresponding minor of the matrix (1.91) also remains different from zero. ∎

Remark 1.55. Non-controllable pairs do not possess the property of structural stability. If the pair (A, B) is not controllable, then there exists a pair (A_1, B_1) of equal dimension, such that the pair $(A + \epsilon A_1, B + \epsilon B_1)$ for arbitrary small $|\epsilon| > 0$ becomes controllable.

8. The vertical pair $[A, C]$ built from the constant $m \times m$ matrix A and $n \times m$ matrix C is called *observable*, if the vertical pair of polynomial matrices $[\lambda I_m - A, \ C]$ is irreducible. Obviously, the pair $[A, C]$ is observable, if and only if the horizontal pair (A', C') is controllable, where the prime means the transposition operation. Due to this reason, observable pairs possess all the properties that have been derived above for controllable pairs. Especially, observable pairs are structural stable.

1.14 Cyclic Matrices

1. The constant $n \times n$ matrix A is said to be *cyclic*, if the assigned characteristic matrix $A_\lambda = \lambda I_n - A$ is simple in the sense of the definition in Section 1.11, see [69, 78, 191].

Cyclic matrices are provided with the important property of structural stability, as is substantiated by the next theorem.

Theorem 1.56. *Let the cyclic $n \times n$ matrix A, and an arbitrary $n \times n$ matrix B be given. Then there exists a positive number $\epsilon_0 > 0$, such that for $|\epsilon| < \epsilon_0$ all matrices $A + \epsilon B$ become cyclic.*

Proof. Let

$$\det(\lambda I_n - A) = d_A(\lambda), \quad \deg d_A(\lambda) = n\,.$$

Then we obtain

$$\det(\lambda I_n - A - \epsilon B) = d_A(\lambda) + \epsilon d_1(\lambda, \epsilon)$$

with $\deg d_1(\lambda, \epsilon) < n$ for all ϵ. Therefore, by virtue of Theorem 1.37, there exists an ϵ_0, such that for $|\epsilon| < \epsilon_0$ the matrix $\lambda I_n - A - \epsilon B$ remains simple, *i.e.* the matrix $A + \epsilon B$ is cyclic. ■

2. Square constant matrices that are not cyclic, will be called in future *composed*. Composed matrices are not equipped with the property of structural stability in the above defined sense. For any composed matrix A, we can find a matrix B, such that the sum $A + \epsilon B$ becomes cyclic, as small even $|\epsilon| > 0$ is chosen. Moreover, the sum $A + \epsilon B$ will become composed only in some special cases. This fact is illustrated by a 2×2 matrix in the next example.

Example 1.57. As follows from Theorem 1.51, any composed 2×2 matrix A is similar to the matrix

$$B = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = aI_2\,, \tag{1.92}$$

where $a =$ const., so we have

$$A = LBL^{-1} = B\,.$$

Therefore, the set of all composed matrices in $\mathbb{C}_{22}$ is determined by Formula (1.92) for any a. Assume now the matrix $Q = A + F$ to be composed. Then

$$Q = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix}$$

is true, and hence

$$F = B - Q = \begin{bmatrix} a - q & 0 \\ 0 & a - q \end{bmatrix}$$

becomes an composed matrix. When the 2×2 matrix A is composed, then the sum $A + F$ still becomes onerous, if and only if the matrix F is composed too. □

3. The property of structural stability of cyclic matrices allows a probability-theoretic interpretation. For instance, the following statement is true:

Let $A \in \mathbb{F}_{nn}$ be a composed matrix and $B \in \mathbb{F}_{nn}$ any random matrix with independent entries that are equally distributed in a certain interval $\alpha \leq b_{ik} \leq \beta$. Then the sum $A+B$ with probability 1 becomes a cyclic matrix.

4. The property of structural stability has great practical importance. Indeed, let for instance the differential equation of a certain linear process be given in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax + Bu\,, \quad A = A_0 + \Delta A\,,$$

where x is the state vector, and A_0, ΔA are constant matrices, where A_0 is cyclic. The matrix ΔA manifests the unavoidable errors during the set up and calculation of the matrix A. From Theorem 1.56 we conclude that the matrix A remains cyclic, if the deviation ΔA satisfies the conditions of Theorem 1.56. If however, the matrix A is composed, then this property can be lost due to the imprecision characterised by the matrix ΔA, as tiny this ever has been with respect to the norm.

5. Assume

$$d(\lambda) = \lambda^n + d_1\lambda^{n-1} + \ldots + d_n \tag{1.93}$$

to be a monic polynomial. Then the $n \times n$ matrix A_F of the form

$$A_F = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -d_n & -d_{n-1} & -d_{n-2} & \ldots & -d_2 & -d_1 \end{bmatrix} \tag{1.94}$$

is called its *accompanying (horizontal) Frobenius matrix* with respect to the polynomial $d(\lambda)$. Moreover, we consider the *vertical accompanying Frobenius matrix*

$$\bar{A}_F = \begin{bmatrix} 0 & 0 & \ldots & 0 & -d_n \\ 1 & 0 & \ldots & 0 & -d_{n-1} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & 0 & -d_2 \\ 0 & 0 & \ldots & 1 & -d_1 \end{bmatrix}. \tag{1.95}$$

The properties of the matrices (1.94) and (1.95) are analogue, so that we could restrict ourself to the investigation of (1.94). The characteristic matrix of A_F has the form

$$\lambda I_n - A_F = \begin{bmatrix} \lambda & -1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & -1 \\ d_n & d_{n-1} & d_{n-2} & \dots & d_2 & \lambda + d_1 \end{bmatrix} . \tag{1.96}$$

Appending Matrix (1.96) with the column $b = \begin{bmatrix} 0 \dots 0 \; 1 \end{bmatrix}'$, we receive the extended matrix

$$\begin{bmatrix} \lambda & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \lambda & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & -1 & 0 \\ d_n & d_{n-1} & d_{n-2} & \dots & d_2 & \lambda + d_1 & 1 \end{bmatrix} .$$

This matrix is alatent, because it has a minor of n-th order that is equal to $(-1)^{n-1}$. Strength to Theorem 1.35, Matrix (1.96) is simple, and therefore, the matrix A_F is cyclic.

6. By direct calculation we recognise

$$\det(\lambda I_n - A_F) = \lambda^n + d_1 \lambda^{n-1} + \dots + d_n = d(\lambda) .$$

According to the properties of simple matrices, we conclude that the whole of invariant polynomials corresponding to the matrix A_F is presented by

$$a_1(\lambda) = a_2(\lambda) = \dots = a_{n-1}(\lambda) = 1, \quad a_n(\lambda) = f(\lambda) .$$

Let A be any cyclic $n \times n$ matrix. Then the accompanying matrix $A_\lambda = \lambda I_n - A$ is simple. Therefore, by applying equivalence transformations, A_λ might be brought into the form

$$\lambda I_n - A = p(\lambda) \operatorname{diag}\{1, 1, \dots, d(\lambda)\} q(\lambda) ,$$

where the matrices $p(\lambda)$, $q(\lambda)$ are unimodular, and $d(\lambda)$ is the characteristic polynomial of the matrix A. From the last equation, we conclude that the set of invariant polynomials of the cyclic matrix A coincides with the set of invariant polynomials of the accompanying Frobenius matrix of its characteristic polynomial $d(\lambda)$. Hereby, the matrices $\lambda I_n - A$ and $\lambda I_n - A_F$ are equivalent, hence the matrices A and A_F are similar, *i.e.*

$$A = L A_F L^{-1}$$

is true with a certain non-singular matrix L. It can be shown that in case of a real matrix A, also the matrix L could be chosen real.

7. As just defined in (1.76), let

$$J_n(a) = \begin{bmatrix} a & 1 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & \dots & 0 & 0 \\ 0 & 0 & a & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}$$

be a Jordan block. The matrix $J_n(a)$ turns out to be cyclic, because the matrix

$$\begin{bmatrix} \lambda - a & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \lambda - a & -1 & \dots & 0 & 0 & 0 \\ 0 & 0 & \lambda - a & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda - a & -1 & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda - a & -1 \end{bmatrix}$$

is alatent.

Let us represent the polynomial (1.93) in the form

$$d(\lambda) = (\lambda - \lambda_1)^{\mu_1} \cdots (\lambda - \lambda_q)^{\mu_q} \,,$$

where all numbers λ_i are different. Consider the matrix

$$J = \operatorname{diag}\{J_{\mu_1}(\lambda_1), \dots, J_{\mu_q}(\lambda_q)\} \tag{1.97}$$

and its accompanying characteristic matrix

$$\lambda I_n - J = \operatorname{diag}\{\lambda I_{\mu_1} - J_{\mu_1}(\lambda_1), \dots, \lambda I_{\mu_q} - J_{\mu_q}(\lambda_q)\} \,, \tag{1.98}$$

where the corresponding diagonal blocks take the shape

$$\lambda I_{\mu_i} - J_{\mu_i}(\lambda_i) = \begin{bmatrix} \lambda - \lambda_i & -1 & 0 & \dots & 0 & 0 \\ 0 & \lambda - \lambda_i & -1 & \dots & 0 & 0 \\ 0 & 0 & \lambda - \lambda_i & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda - \lambda_i & -1 \\ 0 & 0 & 0 & \dots & 0 & \lambda - \lambda_i \end{bmatrix} . \tag{1.99}$$

Obviously, we have

$$\det[\lambda I_{\mu_i} - J_{\mu_i}(\lambda_i)] = (\lambda - \lambda_i)^{\mu_i}$$

so that from (1.98), we obtain

$$\det(\lambda I_n - J) = (\lambda - \lambda_1)^{\mu_1} \cdots (\lambda - \lambda_q)^{\mu_q} = d(\lambda)\,.$$

At the same time, using (1.98) and (1.99), we find

$$\operatorname{rank}(\lambda_i I_n - J) = n - 1, \quad (i = 1, \ldots, q)$$

that means, Matrix (1.98) is cyclic. Therefore, Matrix (1.97) is similar to the accompanying Frobenius matrix of the polynomial (1.93), thus

$$J = L A_F L^{-1}\,,$$

where L in general is a complex non-singular matrix.

1.15 Simple Realisations and Their Structural Stability

1. The triple of matrices $a(\lambda)$, $b(\lambda)$, $c(\lambda)$ of dimensions $p \times p$, $p \times m$, $n \times p$, according to [69] and others, is called a *polynomial matrix description* (PMD)

$$\tau(\lambda) = (a(\lambda), b(\lambda), c(\lambda))\,. \tag{1.100}$$

The integers n, p, m are the dimension of the PMD. In dependence on the membership of the entries of the matrices $a(\lambda)$, $b(\lambda)$, $c(\lambda)$ to the sets $\mathbb{F}[\lambda]$, $\mathbb{R}[\lambda]$, $\mathbb{C}[\lambda]$, the sets of all PMDs with dimension n, p, m are denoted by $\mathbb{F}_{npm}[\lambda]$, $\mathbb{R}_{npm}[\lambda]$, $\mathbb{C}_{npm}[\lambda]$, respectively.

A PMD (1.100) is called *minimal*, if the pairs $(a(\lambda), b(\lambda))$, $[a(\lambda), c(\lambda)]$ are irreducible.

2. A PMD of the form

$$\tau(\lambda) = (\lambda I_p - A, B, C)\,, \tag{1.101}$$

where A, B, C are constant matrices, is said to be an *elementary*. Every elementary PMD (1.101) is characterised by a triple of constant matrices A, $(p \times p)$; B, $(p \times m)$; C, $(n \times p)$. The triple (A, B, C) is called a realisation of the linear process in state space, or shortly *realisation*. The numbers n, p, m are named the *dimension* of the elementary realisation. The set of all realisations with given dimension is denoted by $\mathbb{F}_{npm}$, $\mathbb{R}_{npm}$, $\mathbb{C}_{npm}$, respectively.

Suppose the $p \times p$ matrix Q to be non-singular. Then the realisations (A, B, C) and (QAQ^{-1}, QB, CQ^{-1}) are called *similar*.

3. The realisation (A, B, C) is called *minimal*, if the pair (A, B) is controllable and the pair $[A, C]$ is observable, *i.e.* the elementary PMD (1.101) is minimal. A minimal realisation with a cyclic matrix A is called a *simple realisation*. The set of all minimal realisations of a given dimension will be symbolised by $\bar{\mathbb{F}}_{npm}$, $\bar{\mathbb{R}}_{npm}$, $\bar{\mathbb{C}}_{npm}$ respectively, and the set of all simple realisations by $\mathbb{F}^s_{npm}$, $\mathbb{R}^s_{npm}$ $\mathbb{C}^s_{npm}$. For a simple realisation $(A, B, C) \in \mathbb{R}^s_{npm}$

always exists a similar realisation $(Q_J A Q_J^{-1}, Q_J B, C Q_J^{-1}) \in \mathbb{C}^s_{npm}$, where the matrix $Q_J A Q_J^{-1}$ is of Jordan canonical form. Such a simple realisation is called a *Jordan realisation.* Moreover, for this realisation, there exists a similar realisation $(Q_F A Q_F^{-1}, Q_F B, C Q_F^{-1}) \in \mathbb{R}^s_{npm}$, where the matrix $Q_F A Q_F^{-1}$ is a Frobenius matrix of the form (1.94). Such a simple realisation is called a *Frobenius realisation.*

4. Simple realisations possess the important property of structural stability, as the next theorem states.

Theorem 1.58. *Let the realisation (A, B, C) of dimension n, p, m be simple, and (A_1, B_1, C_1) be an arbitrary realisation of the same dimension. Then there exists an $\epsilon_0 > 0$, such that the realisation $(A + \epsilon A_1, B + \epsilon B_1, C + \epsilon C_1)$ for all $|\epsilon| < \epsilon_0$ remains simple.*

Proof. Since the pair (A, B) is controllable and the pair $[A, C]$ is observable, there exists, owing to Theorem 1.54, an $\epsilon_1 > 0$, such that the pair $(A + \epsilon A_1, B + \epsilon B_1)$ becomes controllable and the pair $[A + \epsilon A_1, C + \epsilon C_1]$ observable for all $|\epsilon| < \epsilon_1$. Furthermore, due to Theorem 1.56, there exists an $\epsilon_2 > 0$, such that the matrix $A + \epsilon A_1$ becomes cyclic for all $|\epsilon| < \epsilon_2$. Consequently, for $|\epsilon| < \min(\epsilon_1, \epsilon_2) = \epsilon_0$ all realisations $(A + \epsilon A_1, B + \epsilon B_1, C + \epsilon C_1)$ are simple. ∎

Remark 1.59. Realisations that are not simple, are not provided by the property of structural stability. For instance, from the above considerations we come to the following conclusion:
Let the realisation (A, B, C) be not simple, and (A_1, B_1, C_1) be a random realisation of equal dimension, where the entries of the matrices $A_1,\ B_1,\ C_1$ are in the whole statistically independent and equally distributed in a certain interval $[\alpha, \beta]$. Then the realisation $(A + A_1, B + B_1, C + C_1)$ will be simple with probability 1.

5. Theorem 1.58 has fundamental importance for developing methods on base of a mathematical description of linear time-invariant multivariable systems. The dynamics of such systems are described in continuous time by state-space equations of the form

$$y = Cx, \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = Ax + Bu\,, \tag{1.102}$$

corresponding to the realisation (A, B, C). In practical investigations, we always will meet $A = A_0 + \Delta A$, $B = B_0 + \Delta B$, $C = C_0 + \Delta C$, where (A_0, B_0, C_0) is the nominal realisation and the realisation $(\Delta A, \Delta B, \Delta C)$ characterises inaccuracies due to finite word length *etc.* Now, if the nominal realisation is simple, then at least for sufficiently small deviations $(\Delta A, \Delta B, \Delta C)$, the simplicity is preserved. Analogue considerations are possible for the description of the dynamics of discrete-time systems, where

$$y_k = Cx_k, \quad x_{k+1} = Ax_k + Bu_k \tag{1.103}$$

is used.

If however, the nominal realisation (A, B, C) is not simple, then the structural properties will, roughly spoken, not be preserved even for tiny deviations.

6. In principle in many cases, we can find suitable bounds of disturbances for which a simple realisation remains simple. For instance, let the matrices $A_1,\ B_1,\ C_1$ depend continuously on a scalar parameter α, such that $A_1 = A_1(\alpha)$, $B_1 = B_1(\alpha)$, $C_1 = C_1(\alpha)$ with $A_1(0) = O_{pp}$, $B_1(0) = O_{pm}$, $C_1(0) = O_{np}$ is valid. Now, if the parameter α increases from zero to positive values, then the realisation $(A+A_1(\alpha), B+B_1(\alpha), C+C_1(\alpha)$ for $0 \leq \alpha < \alpha_0$ remains simple, where α_0 is the smallest positive number, for which at least one of the following conditions takes place:

a) The pair $(A + A_1(\alpha_0), B + B_1(\alpha_0))$ is not controllable.
b) The pair $[A + A_1(\alpha_0), C + C_1(\alpha_0)]$ is not observable.
c) The matrix $A + A_1(\alpha_0)$ is not cyclic.

http://www.springer.com/978-1-84628-431-1

Multivariable Computer-controlled Systems
A Transfer Function Approach
Rosenwasser, E.N.; Lampe, B.
2006, XVIII, 478 p., Hardcover
ISBN: 978-1-84628-431-1