Chapter 5
Tensor Products, Bezoutians, and Stability

Mathematical structures often start from simple ones and are extended by various constructions to structures of increasing complexity. This process is to be controlled, and the guiding lines should include, among other things, applicability to problems of interest. The present chapter is devoted to a circle of ideas from abstract linear algebra that covers several topics of interest to us because of their applicability to the study of linear systems. These topics include bilinear forms defined on vector spaces, module homomorphisms over various rings, and the analysis of classes of special structured matrices such as Bezoutian, Hankel, and Toeplitz matrices. The connections to algebraic methods for analyzing 2D systems, i.e., modules over the ring $\mathbb{F}[z,w]$ of polynomials in two variables, are also explored. The unifying tools are tensor products defined for modules and vector spaces. The interaction between the diverse areas of structured matrix analysis, polynomial modules, and quadratic form theory becomes particularly evident within this context.

Tensor products provide a link between multilinear algebra and classical linear algebra by enabling one to represent multilinear maps as linear functions on a tensor product space. This technique applies both to vector spaces over a field and modules over a ring. The scenario becomes interesting and rich in the context of polynomial and rational models, which are modules over the ring $\mathbb{F}[z]$ and, at the same time, finite-dimensional vector spaces over the field $\mathbb{F}$. In this situation, the result of forming the tensor product depends critically on the algebraic context within which the tensor product is formed, i.e., whether one considers the tensor product over the field $\mathbb{F}$ or over the ring $\mathbb{F}[z]$. Taking the tensor product of two polynomial models $\mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}$ over the field $\mathbb{F}$ yields a space of polynomials in two variables. In contrast, the tensor product $\mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}$ over the ring $\mathbb{F}[z]$ is a module of polynomials in one variable. In this chapter these constructions will be explained and their algebraic properties worked out in detail. As a useful byproduct, explicit characterizations of intertwining maps and homomorphisms of polynomial

P.A. Fuhrmann, U. Helmke, *The Mathematics of Networks of Linear Systems*, Universitext, DOI 10.1007/978-3-319-16646-9_5

models using tensor products will be derived. Following classical duality theory of polynomial models, the mutual relationships between tensor products of polynomial models and spaces of homomorphisms are expressed via the following commutative diagram:

$$\begin{array}{ccc}
(\mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}^*)^* & \xrightarrow{\quad \Phi \quad} & \mathrm{Hom}_{\mathbb{F}}(\mathscr{X}, \mathscr{Y}) \\
\uparrow \iota^* & & \uparrow i \\
(\mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}^*)^* & \xrightarrow{\quad \phi \quad} & \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X}, \mathscr{Y})
\end{array}$$

The horizontal maps denote natural isomorphisms and the vertical maps canonical injections. Although these maps are canonically defined, they crucially involve the construction of dual spaces. Therefore, to obtain a more concrete version of this diagram, a suitable duality theory for power series in two variables is first developed. This will enable us to find explicit descriptions for tensor products and duals of polynomial models and lead us, finally, to the concrete form of the preceding commutative diagram as

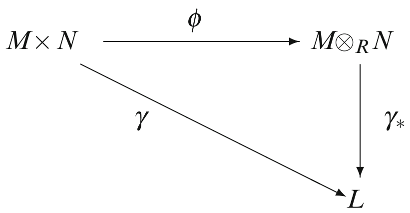
$$\begin{array}{ccc}
X_{D_2(z)\otimes D_1^\top(w)} & \xrightarrow{\quad \Psi \quad} & \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2}) \\
\uparrow \beta & & \uparrow i \\
X^{D_2\otimes I} \cap X^{I\otimes D_1^\top} & \xrightarrow{\quad \psi \quad} & \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})
\end{array}$$

The construction of this commutative diagram is a central purpose of this chapter. It leads to a deeper understanding of the characterization of homomorphisms given in Chapter 3, provides us with a coordinate-free construction of Bezoutian matrices, and yields elegant matrix rank characterizations of coprimeness for matrix polynomials. Historically, the Euclidean algorithm applied to scalar polynomials suggested finding matrix criteria for the coprimeness of two polynomials. This led to the introduction, by Sylvester, of the Bezoutian and resultant matrices. Quadratic (and Hermitian) forms over the field of complex numbers were used efficiently by Hermite (1856) in his work on the root location of polynomials. Generalized Bezoutians, generated by a quadruple of polynomial matrices, originated in the work of Anderson and Jury (1976) in the analysis of coprimeness of polynomial matrices and the solvability of Sylvester-type equations. An interesting advance was the characterization of the Bezoutian matrix as a matrix representation of an intertwining map. Intertwining maps between polynomial models were characterized in Theorem 3.20, providing a powerful tool for the study of multivariable linear systems. Using tensor product representations of module homomorphisms, another proof of this important result will be given. On the way, polynomial versions of classical matrix equations are derived, of which the Sylvester and Lyapunov equations are important special cases. These polynomial Sylvester equations will also prove useful in deriving stability tests for nonsingular matrix polynomials.

5.1 Tensor Products of Modules

Tensor products of modules are at center stage of this chapter, so it is only appropriate to give a working definition. In algebra, the tensor product of two modules is defined abstractly via a universal property as follows.

Definition 5.1. Let R be a commutative ring with identity, and let M,N,L be R-modules. An R-module $M \otimes_R N$ is called a **tensor product** of M and N if there exists an R-bilinear map $\phi : M \times N \longrightarrow M \otimes_R N$ such that for every R-bilinear map $\gamma : M \times N \longrightarrow L$ there exists a unique R-homomorphism $\gamma_* : M \otimes_R N \longrightarrow L$ that makes the following diagram commutative:



Although the notation $M \otimes_R N$ for the tensor product of two modules seems to suggest this, one should beware of assuming that the elements of $M \otimes_R N$ can be represented as single tensor products $m \otimes_R n$ of the elements $m \in M, n \in N$. In fact, the elements of the tensor product are finite sums $\sum_{i=1}^{k} m_i \otimes_R n_i$ and may not simplify to a decomposable representation of the form $m \otimes_R n$. Note further that, according to Definition 5.1, there may be several, necessarily isomorphic, tensor products.

One can give an abstract construction of a tensor product that is briefly sketched as follows; see, for example, Hungerford (1974). Let $< M \times N >$ denote the free R-module of all finite formal linear combinations $\sum_{i=1}^{k} r_i(m_i, n_i)$, where $r_i \in R$, $m_i \in M$, and $n_i \in N$. Let I_R denote the R-submodule of $< M \times N >$ that is generated by elements of the form

1. $r(m,n) - (rm,n),\ r(m,n) - (m,rn)$;
2. $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$;
3. $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$.

Then the quotient R-module

$$M \otimes_R N := < M \times N > / I_R, \tag{5.1}$$

together with the map $\phi : M \times N \longrightarrow M \otimes_R N,\ (m,n) \mapsto m \otimes n := (m,n) + I_R$, satisfies the universal properties of a tensor product and thus serves as a model for the tensor product, unique up to isomorphisms. However, because most of the spaces we use have functional representations, one looks for concrete representations of the various tensor products encountered. As it is, such representations turn out to be amenable to explicit computations.

Listed below are a few basic properties of tensor products. Relative to direct sums, tensor products have the following distributivity and associativity properties:

$$\begin{aligned} (\oplus_{i=1}^{k} M_i) \otimes_R N &\simeq \oplus_{i=1}^{k} (M_i \otimes_R N), \\ M \otimes_R (\oplus_{j=1}^{l} N_j) &\simeq \oplus_{j=1}^{l} (M \otimes_R N_j), \\ M \otimes_R (N \otimes_R P) &\simeq (M \otimes_R N) \otimes_R P. \end{aligned} \tag{5.2}$$

Another useful isomorphism is

$$M \otimes_R N \simeq N \otimes_R M,$$

via the map that sends $m \otimes n$ to $n \otimes m$.

Concerning ring extensions, the following lemma is useful.

Lemma 5.2. *Let S be a subring of R, and let M and N be R-modules and L an S-module. Then:*

1. *The unique S-linear map $M \otimes_S N \longrightarrow M \otimes_R N$, which maps each element $m \otimes_S n$ to $m \otimes_R n$, is surjective.*
2. *Suppose that $b : M \times N \longrightarrow L$ is an S-bilinear map that satisfies*

$$b(rm,n) = b(m,rn)$$

for $r \in R$, $m \in M$, and $n \in N$. Then there exists a unique S-linear map

$$B : M \otimes_R N \longrightarrow L,$$

with $B(m \otimes n) = b(m,n)$.

Proof. Here it pays off to work with the abstract definition of a tensor product. The map $M \times N \longrightarrow M \otimes_R N, (m,n) \mapsto m \otimes_R n$ is S-bilinear and therefore induces a unique S-linear map $M \otimes_S N \longrightarrow M \otimes_R N$ that maps each $m \otimes_S n$ to $m \otimes_R n$. In fact, by inspection of (5.1), one sees that $I_S \subset I_R$ is valid and therefore induces a well-defined surjective map $< M \times N > /I_S \longrightarrow < M \times N > /I_R$, with $(m,n) + I_S \mapsto (m,n) + I_R$. This completes the proof of the first part.

Since each element of $M \otimes_R N$ is of the form $t = m_1 \otimes n_1 + \cdots + m_r \otimes n_r$, the additivity of B implies $B(t) = b(m_1,n_1) + \cdots + b(m_r,n_r)$. This implies the uniqueness of B. To prove the existence of B, define $B(r(m,n)) := b(rm,n)$ for all

$r \in R, m \in M, n \in N$. This extends to a well-defined map $\overline{B}$ on $< M \times N >$. The bilinearity of b implies that $\overline{B}$ is additive. Since $b(rm,n) = b(m,rn)$, it follows that $\overline{B}$ vanishes on I_R and therefore induces a unique S-linear map $B : M \otimes_R N \longrightarrow L$, with $B(m \otimes n) = b(m,n)$. This completes the proof. ■

The tensor product of linear maps is defined as follows. Let $f : M_1 \longrightarrow N_1$ and $g : M_2 \longrightarrow N_2$ be R-linear maps between R-modules M_1, N_1 and M_2, N_2, respectively. Then $f \times g : M_1 \times M_2 \longrightarrow N_1 \otimes_R N_2, (m_1, m_2) \mapsto f(m_1) \otimes f(m_2)$ is R-bilinear and therefore extends in a natural way to an R-linear map,

$$f \otimes g : M_1 \otimes_R M_2 \longrightarrow N_1 \otimes_R N_2,$$

that maps each element $m_1 \otimes m_2$ to $f(m_1) \otimes f(m_2)$. We refer to $f \otimes g$ as the **tensor product** of f with g. It is clear that the tensor product of two module isomorphisms is a module isomorphism. If concrete matrix representations of f and g are given, the matrix representation of the tensor product $f \otimes g$ is equivalent to the so-called Kronecker product of matrices; we will return to this in Section 5.2 when we examine tensor products of polynomial models. Since polynomial models are concrete representations of polynomial quotient modules, it is only natural to expect that the analysis of tensor products of polynomial quotient modules, and of associated module homomorphisms, should prove useful for linear systems theory, and therefore particular attention will be paid to this topic.

The following result provides a very useful identification of tensor products of quotient modules. Let M_1 and M_2 be R-modules, with R a commutative ring. Let $N_i \subset M_i$ be submodules. The quotient spaces M_i/N_i then have a natural R-module structure.

Proposition 5.3. *Let N be the submodule of $M_1 \otimes_R M_2$ defined as*

$$N := N_1 \otimes_R M_2 + M_1 \otimes_R N_2.$$

The R-linear map

$$f : M_1 \otimes_R M_2 \longrightarrow M_1/N_1 \otimes_R M_2/N_2, \quad m_1 \otimes m_2 \mapsto (m_1 + N_1) \otimes (m_2 + N_2)$$

defines the following isomorphism of R-modules:

$$M_1/N_1 \otimes_R M_2/N_2 \simeq (M_1 \otimes_R M_2)/N.$$

Proof. By the construction of f, the kernel of f is contained in N. Moreover, f is surjective. Thus f induces a surjective homomorphism $\overline{f} : (M_1 \otimes_R M_2)/N \longrightarrow M_1/N_1 \otimes_R M_2/N_2$. The map

$$g : M_1/N_1 \otimes_R M_2/N_2 \longrightarrow (M_1 \otimes_R M_2)/N, \quad (m_1 + N_1) \otimes (m_2 + N_2) \mapsto m_1 \otimes m_2 + N$$

is well defined and is a left inverse to $\overline{f}$. Thus $\overline{f}$ is an isomorphism. ■

The next result is useful in establishing isomorphisms between tensor product spaces.

Proposition 5.4. *Let S be a subring of R and $k \geq 3$. Let $M_1, \ldots, M_k$ be R-modules and N an S-module. Suppose that $f : M_1 \times \cdots \times M_k \longrightarrow N$ is an S-multilinear function that satisfies*

$$f(m_1, \ldots, m_{k-2}, rm_{k-1}, m_k) = f(m_1, \ldots, m_{k-2}, m_{k-1}, rm_k)$$

for all $r \in R$ and $m_i \in M_i$. Then there exists a unique S-multilinear function

$$F : M_1 \times \cdots \times M_{k-2} \times (M_{k-1} \otimes_R M_k) \longrightarrow N,$$

with

$$F(m_1, \ldots, m_{k-2}, m_{k-1} \otimes m_k) = f(m_1, \ldots, m_k). \tag{5.3}$$

Proof. Since the elements of $M \otimes_R N$ are finite sums of elementary tensors $a \otimes b$, the uniqueness of F follows from (5.3) and the additivity of F in the last argument. To construct F, we fix $m_1, \ldots, m_{k-2}$. By Lemma 5.2, there exists a unique S-linear map $F_{m_1 \cdots m_{k-2}} : M \otimes_R N \longrightarrow L$, with

$$F_{m_1 \ldots m_{k-2}}(m_{k-1} \otimes m_k) = f(m_1, \ldots, m_{k-2}, m_{k-1}, m_k).$$

By the S-multilinearity of f, this yields the desired S-linear map $F : M_1 \times \cdots \times M_{k-2} \times (M_{k-1} \otimes_R M_k) \longrightarrow N$, satisfying

$$F(m_1, \ldots, m_{k-2}, m_{k-1} \otimes m_k) = F_{m_1 \ldots m_{k-2}}(m_{k-1} \otimes m_k).$$

■

Consider a commutative ring R, and let $S \subset R$ be a subring. Let M and N be R-modules, and let L be an S-module. Note that the space $\mathrm{Hom}_S(N, L)$ of S-linear maps becomes an R-module by defining

$$(r \cdot f)(n) = f(rn)$$

for all $r \in R, n \in N$ and $f \in \mathrm{Hom}_S(N, L)$. For greater generality, one denotes by $\mathrm{Bil}_{S,R}(M, N; L)$ the set of all S-bilinear maps $f : M \times N \longrightarrow L$ that satisfy

$$f(rm, n) = f(m, rn) \tag{5.4}$$

for all $(m, n) \in M \times N$ and $r \in R$. In the special case where $S = R$, one uses the simplified notation $\mathrm{Bil}_R(M, N; L)$. It is easily seen that $\mathrm{Bil}_{S,R}(M, N; L)$ is, in a natural way, an R-module with respect to the R-scalar product

$$(r \cdot f)(m, n) = f(rm, n), \quad r \in R.$$

Proposition 5.5. *Let R be a commutative ring, let $S \subset R$ be a subring and let M and N be R-modules and L an S-module. Then the following are R-module isomorphisms:*

$$\mathrm{Hom}_S(M \otimes_R N, L) \simeq \mathrm{Bil}_{S,R}(M, N; L) \simeq \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, L)).$$

Proof. Let $f \in Bil_{S,R}(M,N;L)$, and let $m \in M$. Defining, for $n \in N$, $f_m(n) = f(m,n)$, it follows that $f_m : M \longrightarrow \mathrm{Hom}_S(N,L)$ is S-linear for all $m \in M$. Since f satisfies (5.4), then $(r \cdot f_m)(n) = f(m,rn) = f(rm,n) = f_{rm}(n)$ for all $r \in R$. Thus $f_{rm} = r \cdot f_m$, i.e., the map f_m belongs to $\mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L))$. On the other hand, given $g \in \mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L))$, a map $f : M \times N \longrightarrow L$ is defined by $f(x,y) = g(x)(y)$, which is necessarily S-bilinear and satisfies (5.4) for each $r \in R$. This proves the isomorphism $\mathrm{Bil}_{S,R}(M,N;L) \simeq \mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L))$. To prove the existence of an R-linear isomorphism

$$\mathrm{Hom}_S(M \otimes_R N, L) \simeq \mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L)),$$

consider the map $\Theta : \mathrm{Hom}_S(M \otimes_R N, L) \longrightarrow \mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L)), \Theta f = g$, where

$$\Theta(f)(m)(n) := f(m \otimes_R n), \qquad f \in \mathrm{Hom}_S(M \otimes_R N, L).$$

Clearly, Θ is R-linear. Similarly, the map

$$\Psi : \mathrm{Hom}_R(M, \mathrm{Hom}_S(N,L)) \longrightarrow \mathrm{Hom}_S(M \otimes_R N, L),$$

which is defined by

$$\Psi g(m \otimes_R n) := g(m)(n),$$

is R-linear and $\Psi \circ \Theta$ and $\Theta \circ \Psi$ are identity maps. The result follows. ∎

The duality properties of tensor products are of interest to us. Recall that the **algebraic dual module** M' of an R-module M is defined by

$$M' = \mathrm{Hom}_R(M, R), \tag{5.5}$$

i.e., by the space of all R-linear functionals on M or, equivalently, by the space of all R-homomorphisms of M into R. For the special case $L = R$, Proposition 5.5 implies the module isomorphisms

$$(M \otimes_R N)' = \mathrm{Hom}_R(M \otimes_R N, R) \simeq \mathrm{Hom}_R(M, N'). \tag{5.6}$$

Clearly, if M is an R-torsion module and the ring R has no zero divisors, then $M' = 0$. Unfortunately, the definition of the algebraic dual in the context of modules, namely, by (5.5), is of little use for the applications we have in mind. Of course, given a subring $S \subset R$, one can replace the algebraic dual M' with the S-dual

$$M^* := \mathrm{Hom}_S(M,S).$$

This is still an R-module that has in general better duality properties than M'. The best choice here would be to take S as a subfield of R, and this is what will be done in the sequel. The objects of interest to us are polynomial and rational models. Both have two structures: they are vector spaces over the field $\mathbb{F}$ and modules over the polynomial ring $\mathbb{F}[z]$. As $\mathbb{F}[z]$-modules, they are finitely generated torsion modules, and hence their algebraic dual, defined by (5.5), is the zero module. In much the same way, all objects defined in the subsequent isomorphism (5.6) are trivial. To overcome this problem, two module structures, over $\mathbb{F}$ and $\mathbb{F}[z]$ respectively, will be used and the algebraic dual replaced by the vector space dual.

1. Tensor products of vector spaces

Tensor products of vector spaces over a field $\mathbb{F}$ are studied next in somewhat more detail. This is of course a much simpler situation in which most of the pathologies encountered in studying tensor products over a ring disappear. As indicated previously, one needs to introduce vector space duality. For a finite-dimensional $\mathbb{F}$-vector space $\mathscr{X}$, the **vector space dual** is defined by

$$\mathscr{X}^* = \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathbb{F}).$$

The **annihilator** of a subspace $\mathscr{V} \subset \mathscr{X}$ is defined as the linear subspace

$$\mathscr{V}^{\perp} := \{\lambda \in \mathscr{X}^* \mid \lambda|_{\mathscr{V}} = 0\}.$$

If $\mathscr{X}$ is finite-dimensional, then so is $\mathscr{V}^{\perp}$ and

$$\dim \mathscr{V}^{\perp} = \dim \mathscr{X} - \dim \mathscr{V}.$$

Finite-dimensional vector spaces are reflexive, i.e., the isomorphism $\mathscr{X}^{**} \simeq \mathscr{X}$ is satisfied. In fact, these two spaces can be identified by letting each vector $x \in \mathscr{X}$ act on an element $x^* \in \mathscr{X}^*$ by $x(x^*) = x^*(x)$. We now take a closer look at the case of tensor products of two finite-dimensional $\mathbb{F}$-linear spaces $\mathscr{X},\mathscr{Y}$. Let $\mathscr{B}_X = \{f_i\}_{i=1}^n$, $\mathscr{B}_Y = \{g_i\}_{i=1}^m$ be bases of $\mathscr{X}$ and $\mathscr{Y}$, respectively. Let $\mathscr{B}_X^* = \{\phi_i\}_{i=1}^n$ be the basis of $\mathscr{X}^*$, which is dual to $\mathscr{B}_X$, i.e., it satisfies $\phi_i(f_j) = \delta_{ij}$. For a linear transformation $T \in \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y})$, let $t_{ij} \in \mathbb{F}$ be defined by

$$Tf_j = \sum_{i=1}^{m} t_{ij} g_i, \qquad j = 1,\ldots,n. \tag{5.7}$$

Thus $[T]_{\mathscr{B}_X}^{\mathscr{B}_Y} = (t_{ij})$ is the matrix representation with respect to this pair of bases. On the other hand, we consider the tensor product $\mathscr{Y} \otimes \mathscr{X}^*$, which is generated by the basis elements $g_i \otimes \phi_k$. Associate with $g_i \otimes \phi_k$ the linear map from $\mathscr{X}$ to $\mathscr{Y}$, defined for $x \in \mathscr{X}$ by

$$(g_i \otimes \phi_k)x = \phi_k(x)g_i.$$

We claim that $\{g_i \otimes \phi_k | i = 1,\dots,m, k = 1,\dots,n\}$ is a basis for $L(\mathscr{X},\mathscr{Y})$. Indeed, $T = \sum_{i=1}^m \sum_{k=1}^n c_{ik} g_i \otimes \phi_k$ implies

$$\begin{aligned} Tf_j &= \sum_{i=1}^m \sum_{k=1}^n c_{ik}(g_i \otimes \phi_k)f_j = \sum_{i=1}^m \sum_{k=1}^n c_{ik}\phi_k(f_j)g_i \\ &= \sum_{i=1}^m \sum_{k=1}^n c_{ik}\delta_{kj}g_i = \sum_{i=1}^m c_{ij}g_i. \end{aligned}$$

Comparing this with (5.7), it follows that $c_{ij} = t_{ij}$. Hence,

$$T = \sum_{i=1}^m \sum_{j=1}^n t_{ij} g_i \otimes \phi_j, \tag{5.8}$$

i.e., $\{g_i \otimes \phi_j\}$ is a basis for $\mathscr{Y} \otimes_{\mathbb{F}} \mathscr{X}^*$. This leads us back to the isomorphism

$$\mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y}) \simeq \mathscr{Y} \otimes_{\mathbb{F}} \mathscr{X}^* \tag{5.9}$$

given by (5.8). The representation (5.8) of $T \in \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y})$ can be simplified. If $\mathrm{rank}\, T = k$, then there exists a minimal-length representation

$$T = \sum_{i=1}^k \psi_i \otimes \phi_i,$$

where $\{\phi_i\}_{i=1}^k$ is a basis of $(\mathrm{Ker}\, T)^\perp \subset \mathscr{X}^*$ and $\{\psi_i\}_{i=1}^k$ is a basis of $\mathrm{Im}\, T \subset \mathscr{Y}$. For a linear transformation $T \in \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y})$, the **adjoint transformation** $T^* \in \mathrm{Hom}_{\mathbb{F}}(\mathscr{Y}^*,\mathscr{X}^*)$ is defined by

$$(T^*y^*)x = y^*(Tx).$$

The map $T \mapsto T^*$ yields the isomorphism

$$\mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y}) \simeq \mathrm{Hom}_{\mathbb{F}}(\mathscr{Y}^*,\mathscr{X}^*).$$

Clearly, (5.9) is a special case of (5.6). Another consequence of (5.9) is the following dimension formula:

$$\dim(\mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}) = \dim \mathscr{X} \cdot \dim \mathscr{Y}. \tag{5.10}$$

Note that in view of the isomorphism (5.9), for $T \in \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y})$ there are two possible matrix representations that turn out to be equal, namely,

$$[T]^{\mathscr{B}_Y}_{\mathscr{B}_X} = [T]^{\mathscr{B}_Y \otimes \mathscr{B}_X^*}.$$

Here $\mathscr{B}_Y \otimes \mathscr{B}_X^*$ is the tensor product of the basis $\mathscr{B}_Y$ of $\mathscr{Y}$ and the basis $\mathscr{B}_X^*$ of X^*, which is dual to the basis $\mathscr{B}_X$ of X.

2. Tensor products of $\mathbb{F}[z]$-modules

Next, the situation of principal interest to us, namely, the case where the vector spaces $\mathscr{X}$ and $\mathscr{Y}$ are actually $\mathbb{F}[z]$-modules, is addressed. The module structure on $\mathscr{X}$ defines a canonical $\mathbb{F}[z]$-module structure on the dual vector space $\mathscr{X}^*$ via

$$(z\cdot\lambda)(x) = \lambda(z\cdot x)$$

for $x \in \mathscr{X}$ and $\lambda \in \mathscr{X}^*$. As a consequence of the reflexivity property, there exists an isomorphism $\mathscr{X} \simeq \mathscr{X}^{**}$ as $\mathbb{F}[z]$-modules. Of particular interest is the establishment of a relation between the operations of forming the tensor product $\mathscr{X} \otimes \mathscr{Y}$ and the space of homomorphisms $\mathrm{Hom}(\mathscr{X},\mathscr{Y})$. The dimension formula (5.10) implies that there is a well-defined $\mathbb{F}$-linear isomorphism

$$\mathscr{Y} \otimes_{\mathbb{F}} \mathscr{X}^* \longrightarrow \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y}), \quad (y,\lambda) \mapsto (x \mapsto \lambda(x)y) \tag{5.11}$$

as long as $\mathscr{X}$ and $\mathscr{Y}$ are finite-dimensional $\mathbb{F}$-vector spaces. It is desirable to establish a similar $\mathbb{F}[z]$-linear isomorphism of modules $\mathscr{Y} \otimes_{\mathbb{F}[z]} \mathscr{X}^* \longrightarrow \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X},\mathscr{Y})$. For a ring R and R-modules M and N, it is well known, see Hilton and Wu (1974), that if M is finitely generated and projective, then the canonical map

$$\phi : N \otimes_R M' \longrightarrow \mathrm{Hom}_R(M,N), \quad (n,\lambda) \mapsto (m \mapsto \lambda(m)n) \tag{5.12}$$

is a module isomorphism. Since every free module is projective, this covers the vector space case. Thus, it seems to be the appropriate generalization of (5.11) to $\mathbb{F}[z]$-modules. Unfortunately, the isomorphism (5.12) excludes the case of M being a torsion module, which is the case of main interest for us. This shows that the isomorphism (5.12) with the algebraic dual M' instead of the vector space dual M^* is of no further use for us.

Let $D(z) \in \mathbb{F}[z]^{m\times m}$ be a nonsingular polynomial matrix, and let X_D denote the associated polynomial model. Standard duality theory for polynomial models, as summarized in Theorem 3.38, shows that there exists an isomorphism of $\mathbb{F}[z]$-modules

$$X_D \simeq X_{D^\top}^* \simeq X_{D^\top}.$$

In fact, the isomorphism $X_D \simeq X_{D^\top}$ follows from the fact that both polynomial matrices $D(z)$ and $D^\top(z)$ have the same invariant factors. The isomorphism

$X_D \simeq X^*_{D^\top}$ follows from Theorem 3.38. The next result explores in more detail the various isomorphisms between spaces of homomorphisms and tensor products.

Proposition 5.6. *Let $\mathbb{F}$ be a field and $\mathscr{X}$ and $\mathscr{Y}$ be $\mathbb{F}[z]$-modules that are finite-dimensional as vector spaces over $\mathbb{F}$. The following assertions are true:*

1. There is a natural $\mathbb{F}[z]$-module isomorphism

$$\mathscr{X} \simeq \mathscr{X}^{**}.$$

2. There is a natural $\mathbb{F}[z]$-module isomorphism

$$\mathrm{Bil}_{\mathbb{F},\mathbb{F}[z]}(\mathscr{X},\mathscr{Y};\mathbb{F}) \simeq (\mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}^*)^* \simeq \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X},\mathscr{Y}). \tag{5.13}$$

3. There are natural $\mathbb{F}$-vector space isomorphisms

$$\mathscr{X}^* \otimes_{\mathbb{F}} \mathscr{Y} \simeq \mathrm{Bil}_{\mathbb{F}}(\mathscr{X},\mathscr{Y};\mathbb{F}) \simeq (\mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}^*)^* \simeq \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y}). \tag{5.14}$$

Proof. 1. For $x \in \mathscr{X}$ let $\varepsilon_x : \mathscr{X} \longrightarrow \mathbb{F}$ denote the linear functional defined as $\varepsilon_x(\lambda) = \lambda(x)$. By the finite dimensionality of $\mathscr{X}$, the natural map $\varepsilon : \mathscr{X} \longrightarrow \mathscr{X}^{**}, x \mapsto \varepsilon_x$ is an $\mathbb{F}$-linear isomorphism. The $\mathbb{F}[z]$-module structure on $\mathscr{X}^{**}$ is defined as $(z \cdot \varepsilon_x)(\lambda) := \varepsilon_x(z\dot{\lambda}) = \lambda(zx)$ for $x \in \mathscr{X}, \lambda \in \mathscr{X}^*$. Therefore, $z \cdot \varepsilon_x = \varepsilon_{zx}$ for all $x \in \mathscr{X}$, which proves that ε is $\mathbb{F}[z]$-linear. The result follows.

2. The isomorphism $(\mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}^*)^* \simeq \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X},\mathscr{Y})$ follows by applying Proposition 5.5 to the case of $S = \mathbb{F}$, $R = \mathbb{F}[z]$, $M = \mathscr{X}$, and $N = \mathscr{Y}^*$, noting that the $\mathbb{F}[z]$-modules $\mathscr{X}^{**}$ and $\mathscr{X}$ are isomorphic.

3. The last two natural isomorphisms in (5.14) follow from Proposition 5.5 by taking $R = S = \mathbb{F}$. The natural isomorphism between $\mathscr{X}^* \otimes_{\mathbb{F}} \mathscr{Y}$ and $\mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y})$ can be constructed explicitly by mapping the generating elements $x^* \otimes_{\mathbb{F}} y$ of the tensor product $\mathscr{X}^* \otimes_{\mathbb{F}} \mathscr{Y}$ onto the linear transformation from $\mathscr{X}$ to $\mathscr{Y}$, defined by $(x^* \otimes_{\mathbb{F}} y)(x) = (x^*(x))y$. ■

Proposition 5.6 implies the following commutative diagram of linear maps:

$$\begin{array}{ccc}
(\mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}^*)^* & \xrightarrow{\Psi} & \mathrm{Hom}_{\mathbb{F}}(\mathscr{X},\mathscr{Y}) \\
\uparrow \iota^* & & \uparrow i \\
(\mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}^*)^* & \xrightarrow{\psi} & \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X},\mathscr{Y})
\end{array}$$

Here ι^* denotes the adjoint transformation of the canonical surjective $\mathbb{F}$-linear map $\iota : \mathscr{X} \otimes_{\mathbb{F}} \mathscr{Y}^* \longrightarrow \mathscr{X} \otimes_{\mathbb{F}[z]} \mathscr{Y}^*$ and therefore is injective. Map i is the canonical inclusion map and ψ and Ψ are the canonical $\mathbb{F}[z]$-linear and $\mathbb{F}$-linear isomorphisms defined by (5.13) and (5.14), respectively. The preceding diagram has the advantage of involving only canonical constructions of polynomial models. Its disadvantage lies in the usage of duality, which makes it hard to write down the maps in concrete form. Therefore, in the sequel, we will aim at a more concrete representation of the spaces appearing in this diagram that does not involve dual spaces. This new diagram will be derived in the following sections and will require concrete representations for tensor product spaces. It will also play an important role in our analysis of Bezoutians.

Next, one extends the duality theory from vector spaces to $\mathbb{F}[z]$-modules.

Definition 5.7. Let M be a module over the ring of polynomials $\mathbb{F}[z]$. The **vector space annihilator** of an $\mathbb{F}[z]$-submodule $N \subset M$ is the submodule of M^* defined as

$$N^\perp = \{\phi \in M^* \mid \phi|_N = 0\}.$$

Using this notation, the dual of a quotient module over $\mathbb{F}[z]$ has a nice representation. In fact, one has the $\mathbb{F}[z]$-linear isomorphism

$$(M/N)^* \simeq N^\perp.$$

The isomorphism (5.6) is extended to the tensor product of two quotient modules.

Proposition 5.8. *Let $N_i \subset M_i$, $i = 1,2$, be $\mathbb{F}[z]$-modules. There is an $\mathbb{F}[z]$-linear isomorphism*

$$\begin{aligned}(M_1/N_1 \otimes_{\mathbb{F}[z]} M_2/N_2)^* &\simeq (M_1 \otimes_{\mathbb{F}[z]} N_2 + N_1 \otimes_{\mathbb{F}[z]} M_2)^\perp \\ &= (N_1 \otimes_{\mathbb{F}[z]} M_2)^\perp \cap (M_1 \otimes_{\mathbb{F}[z]} N_2)^\perp.\end{aligned}$$

Proof. Let N denote the submodule in $M_1 \otimes_{\mathbb{F}[z]} M_2$ that is generated by the spaces $N_1 \otimes_{\mathbb{F}[z]} M_2$ and $M_1 \otimes_{\mathbb{F}[z]} N_2$. Clearly, the equality of annihilators $N^\perp = (N_1 \otimes_{\mathbb{F}[z]} M_2 + M_1 \otimes_{\mathbb{F}[z]} N_2)^\perp$ is true. Thus Proposition 5.3 implies the module isomorphism

$$\begin{aligned}(M_1/N_1 \otimes_{\mathbb{F}[z]} M_2/N_2)^* &\simeq (M_1 \otimes_{\mathbb{F}[z]} M_2)/N \simeq N^\perp = (N_1 \otimes_{\mathbb{F}[z]} M_2 + M_1 \otimes_{\mathbb{F}[z]} N_2)^\perp \\ &= (N_1 \otimes_{\mathbb{F}[z]} M_2)^\perp \cap (M_1 \otimes_{\mathbb{F}[z]} N_2)^\perp.\end{aligned}$$ ■

3. Tensor product spaces of Laurent series

The ambient space for the algebraic analysis of discrete-time linear systems is $\mathbb{F}((z^{-1}))^m$. Thus as a first step one considers the tensor product of such spaces, both taken over the field $\mathbb{F}$ as well as over the ring of polynomials $\mathbb{F}[z]$. Clearly, the polynomial ring $\mathbb{F}[z]$ is a rank one module over itself but an infinite-dimensional vector space over $\mathbb{F}$.

Proposition 5.9. *The following $\mathbb{F}$-linear and $\mathbb{F}[z]$-linear isomorphisms are valid:*

$$\begin{aligned} \mathbb{F}[z]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m &\simeq \mathbb{F}[z,w]^{p\times m}, \quad f\otimes_{\mathbb{F}} g \mapsto f(z)g(w)^\top, \\ \mathbb{F}[z]^p \otimes_{\mathbb{F}[z]} \mathbb{F}[z]^m &\simeq \mathbb{F}[z]^{p\times m}, \quad f\otimes_{\mathbb{F}[z]} g \mapsto f(z)g(z)^\top. \end{aligned} \tag{5.15}$$

Proof. To prove (5.15), one notes that the map $\gamma : \mathbb{F}[z]^p \times \mathbb{F}[z]^m \longrightarrow \mathbb{F}[z,w]^{p\times m}$ that maps a pair of polynomials (f,g) to the polynomial matrix in two variables $f(z)g(w)^\top$ is $\mathbb{F}$-linear and therefore determines a unique $\mathbb{F}$-linear map $\gamma_* : \mathbb{F}[z]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m \longrightarrow \mathbb{F}[z,w]^{p\times m}$ for which $\gamma = \gamma_*\phi$. γ_* is surjective because every element $Q(z,w) \in \mathbb{F}[z,w]^{p\times m}$ is a finite sum $Q(z,w) = \sum_{i=1}^q f_i(z)g_i(w)^\top = \gamma_* \sum_{i=1}^q f_i \otimes_{\mathbb{F}} g_i$. To prove the injectivity of γ_*, we note that $\mathbb{F}[z]^m$ has a basis $\{z^i e_j | i \in \mathbb{Z}_+, j = 1,\ldots,m\}$. Therefore, each element of $\mathbb{F}[z]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m$ in the kernel of γ_* has the form $\xi = \sum_{(i,j)\in I} f_{ij} z^i e_j$, with $\gamma_*(\xi) = \sum_{(i,j)\in I} f_{ij}(z) w^i (e_j)^\top = 0$. Hence, $\sum_{(i,j)\in I} f_{ij}(z)w^i = 0$ for all j, implying that $f_{ij}(z) = 0$ for all i,j. Thus $\xi = 0$. Mutatis mutandis, using the $\mathbb{F}[z]$-bilinear map $\gamma : \mathbb{F}[z]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m \longrightarrow \mathbb{F}[z]^{p\times m}$ that maps a pair of polynomials (f,g) to the polynomial matrix $f(z)g(z)^\top \in \mathbb{F}[z]^{p\times m}$ exhibits an induced $\mathbb{F}[z]$-isomorphism $\mathbb{F}[z]^p \times \mathbb{F}[z]^m \longrightarrow \mathbb{F}[z]^{p\times m}$, which proves the second isomorphism in (5.15). ∎

The surjectivity of (5.15) can be reformulated as follows.

Proposition 5.10. *Every $Q(z,w) \in \mathbb{F}[z,w]^{p\times m}$ has a representation of the form*

$$Q(z,w) = \sum_{i=1}^k R_i(z) P_i^\top(w),$$

with $R_i(z) \in \mathbb{F}[z]^p$ and $P_i(w) \in \mathbb{F}[w]^m$. This implies a factorization

$$Q(z,w) = R(z)P^\top(w),$$

with $R(z) \in \mathbb{F}[z]^{p\times k}$ and $P(w) \in \mathbb{F}[w]^{m\times k}$.

To extend the previous results to Laurent series, several more spaces will be needed. Because the field $\mathbb{F}((z^{-1}))$ of truncated Laurent series has two module structures of interest, namely, with respect to the fields $\mathbb{F}$ and $\mathbb{F}((z^{-1}))$, there are two different tensor products, given by

$$\mathbb{F}((z^{-1}))^p \otimes_{\mathbb{F}} \mathbb{F}((z^{-1}))^m \simeq \mathbb{F}_{\mathrm{sep}}((z^{-1},w^{-1}))^{p\times m} \tag{5.16}$$

and

$$\mathbb{F}((z^{-1}))^p \otimes_{\mathbb{F}((z^{-1}))} \mathbb{F}((z^{-1}))^m \simeq \mathbb{F}((z^{-1}))^{p\times m}. \tag{5.17}$$

These are the analogs of equation (5.15).

Here, $\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))$ denotes the ring of separable truncated Laurent series in the variables z and w, which are of the form $F(z,w) = \sum_{i=1}^{N} f_i(z)g_i(w)$ for finitely many $f_1,\ldots,f_N \in \mathbb{F}((z^{-1}))$, $g_1,\ldots,g_N \in \mathbb{F}((w^{-1}))$. Thus $\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))$ is a proper subset of $\mathbb{F}((z^{-1},w^{-1}))$, the field of truncated Laurent series. By $\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m}$ we denote the module of all $p\times m$ matrices with entries in $\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))$. Rational elements $H(z,w) \in \mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m}$ have representations of the form $H(z,w) = \sum_{i=1}^{k} f_i(z)g_i(w)^\top$, with both $f_i(z)$ and $g_i(z)$ rational. This implies a representation of the form

$$H(z,w) = d(z)^{-1}Q(z,w)e(w)^{-1},$$

with $Q(z,w) \in \mathbb{F}[z,w]^{p\times m}$ and $e(w),d(z)$ nonzero, scalar polynomials.

The isomorphism $\mathbb{F}((z^{-1},w^{-1}))^{p\times m} \simeq \mathbb{F}^{p\times m}((z^{-1},w^{-1}))$ will be routinely used, and we will actually identify the two spaces. The identification $\mathbb{F}^{p\times m}[z,w] = \mathbb{F}[z,w]^{p\times m}$ is a special case. By $\mathbb{F}[z,w]$ we denote the ring of polynomials in the variables z and w and by $\mathbb{F}[[z^{-1},w^{-1}]]$ the ring of formal power series in z^{-1} and w^{-1}. Denote by $\mathbb{F}[z,w]^{p\times m}$ the space of $p\times m$ polynomial matrices. The elements of $z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$ are called *strongly strictly proper*. It should be emphasized that not every strictly proper rational function in two variables belongs to $z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]w^{-1}$. For example, $\frac{1}{z-w} \notin z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]w^{-1}$. It will be convenient to use $\mathbb{F}[[z^{-1},w]^{p\times m}$ to denote the subspace of $\mathbb{F}((z^{-1},w^{-1}))^{p\times m}$ of matrices whose entries are formal power series in z^{-1} and polynomial in w. Thus the elements of $\mathbb{F}[[z^{-1},w]^{p\times m}$ are of the form $F(z,w) = \sum_{i=0}^{N} F_i(z)w^i$ for suitable $F_i(z) \in \mathbb{F}((z^{-1}))^{p\times m}$. Thus $\mathbb{F}[[z^{-1},w]^{p\times m} \subset \mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m}$. The space $\mathbb{F}[z,w^{-1}]]^{p\times m}$ is similarly defined. In the same vein, $\mathbb{F}_{\text{sep}}[[z^{-1},w^{-1}]]^{p\times m}$ denotes the space of $p\times m$ matrix functions of separable formal power series $\sum_{i=1}^{N} f_i(z)g_i(w)^\top$, with $f_i(z) \in \mathbb{F}[[z^{-1}]]^p$, $g_i(z) \in \mathbb{F}[[w^{-1}]]^m$.

The definition of $\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m}$ implies the isomorphism

$$\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m} \simeq \mathbb{F}((z^{-1}))^p \otimes_{\mathbb{F}} \mathbb{F}((z^{-1}))^m.$$

Taking into account the direct sum representations (3.1) as well as (5.2), one computes

$$\begin{aligned}
&\mathbb{F}_{\text{sep}}((z^{-1},w^{-1}))^{p\times m} \simeq (\mathbb{F}[z]^p \oplus z^{-1}\mathbb{F}[[z^{-1}]]^p) \otimes_{\mathbb{F}} (\mathbb{F}[z]^m \oplus z^{-1}\mathbb{F}[[z^{-1}]]^m)\\
&\simeq (\mathbb{F}[z]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m) \oplus (\mathbb{F}[z]^p \otimes_{\mathbb{F}} z^{-1}\mathbb{F}[[z^{-1}]]^m)\\
&\oplus (z^{-1}\mathbb{F}[[z^{-1}]]^p \otimes_{\mathbb{F}} \mathbb{F}[z]^m) \oplus (z^{-1}\mathbb{F}[[z^{-1}]]^p \otimes_{\mathbb{F}} z^{-1}\mathbb{F}[[z^{-1}]]^m)\\
&\simeq \mathbb{F}[z,w]^{p\times m} \oplus \mathbb{F}[z,w^{-1}]]^{p\times m}w^{-1} \oplus z^{-1}\mathbb{F}[[z^{-1},w]^{p\times m} \oplus z^{-1}\mathbb{F}_{\text{sep}}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}.
\end{aligned}$$

To these direct sum representations correspond, respectively, the following projection identities:

$$\begin{aligned} I &= \pi_+^z \otimes I + \pi_-^z \otimes I = I \otimes \pi_+^w + I \otimes \pi_-^w \\ &= \pi_+^z \otimes \pi_+^w + \pi_-^z \otimes \pi_+^w + \pi_+^z \otimes \pi_-^w + \pi_-^z \otimes \pi_-^w. \end{aligned}$$

Our next target is the extension of duality theory to the context of polynomial spaces in two variables. To this end, the space of matrix truncated Laurent series in two variables, i.e.,

$$\mathbb{F}((z^{-1}, w^{-1}))^{p\times m} = \{G(z,w) = \sum_{i=-\infty}^{n_1} \sum_{j=-\infty}^{n_2} G_{ij} z^i w^j\},$$

is introduced. For $G(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$, its **residue** is defined as the coefficient of $z^{-1}w^{-1}$, i.e., $G_{-1,-1}$. In analogy with (3.41), for $G(z,w), H(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$, we define a bilinear form on $\mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$ by

$$\begin{aligned} &= \mathrm{Trace}(H^\top G)_{-1,-1} = \mathrm{Trace} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} H^\top_{-i-1,-j-1} G_{ij} \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \mathrm{Trace}\, H^\top_{-i-1,-j-1} G_{ij}. \end{aligned} \tag{5.18}$$

Note that the sum defining $[G,H]$ contains only a finite number of nonzero terms. Clearly, the form defined in (5.18) is nondegenerate. If $G(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{q\times m}$, $A(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{p\times q}$, and $H(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$, then $A(z,w)G(z,w) \in \mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$ and

$$[AG, H] = [G, A^\top H].$$

It is easy to see that, with respect to the bilinear form (5.18), one has

$$(\mathbb{F}[z,w]^{p\times m})^\perp = \mathbb{F}[z, w^{-1}]]^{p\times m} + \mathbb{F}[[z^{-1}, w]^{p\times m}. \tag{5.19}$$

The next result gives a concrete representation of the dual space of $\mathbb{F}[z,w]^{p\times m}$ that is an extension of Theorem 3.38.

Proposition 5.11. *The vector space dual of* $\mathbb{F}[z,w]^{p\times m}$ *is* $\mathbb{F}$*-linear isomorphic to the space* $(z^{-1}\mathbb{F}[[z^{-1}, w^{-1}]]w^{-1})^{p\times m}$*, i.e.,*

$$(\mathbb{F}[z,w]^{p\times m})^* \simeq (z^{-1}\mathbb{F}[[z^{-1}, w^{-1}]]w^{-1})^{p\times m}.$$

Proof. Clearly, for $H(z,w) \in (z^{-1}\mathbb{F}[[z^{-1}, w^{-1}]]w^{-1})^{p\times m}$ and $Q(z,w) \in \mathbb{F}[z,w]^{p\times m}$, the map $\Phi : \mathbb{F}[z,w]^{p\times m} \longrightarrow \mathbb{F}$, defined by $\Phi(Q) = [Q,H]$, is a linear functional on $\mathbb{F}[z,w]^{p\times m}$. Conversely, suppose Φ is a linear functional on $\mathbb{F}[z,w]^{p\times m}$. For all

$i, j \geq 0$, Φ induces linear functionals Φ_{ij} on $\mathbb{F}^{p\times m}$ by defining $\Phi_{ij}(A) = \Phi(z^i A w^j)$. Every functional Φ_{ij} on $\mathbb{F}^{p\times m}$ has a representation of the form $\Phi_{ij}(A) = \mathrm{Trace}\,(H_{ij}^\top A)$ for a unique $H_{ij} \in \mathbb{F}^{p\times m}$. Defining

$$H(z,w) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} H_{ij} z^{-i-1} w^{-j-1} \in z^{-1}\mathbb{F}[[z^{-1}, w^{-1}]]^{p\times m} w^{-1},$$

it follows that $\Phi(Q) = [Q, H]$. ∎

5.2 Tensored Polynomial and Rational Models

Turning now to a detailed study of tensor products of polynomial and rational models, taken together with duality theory, will enable us to construct, in concrete terms, an isomorphism

$$\mathscr{X}^* \otimes_{\mathbb{F}[z]} \mathscr{Y} \longrightarrow \mathrm{Hom}_{\mathbb{F}[z]}(\mathscr{X}, \mathscr{Y})$$

and, in the process, develop a coordinate-free approach to Bezoutians.

Spaces like $\mathbb{F}[z]^m$ or, more importantly for our purposes, quotient spaces like $\mathbb{F}[z]^m / D(z)\mathbb{F}[z]^m$ have module structures with respect to both the field $\mathbb{F}$, i.e., vector space structures, and the ring of polynomials $\mathbb{F}[z]$. With respect to the characterization of tensor products, the underlying ring is of utmost importance because the tensor product depends very much on the ring used. These two constructs do not exhaust the possibilities, especially where polynomial models are concerned, and we will also study polynomial models defined by the Kronecker product of polynomial matrices. In analyzing the tensor products of two polynomial models, our first objective will be to find concrete representations of the various tensor products. Furthermore, it will be shown that the class of polynomial models is closed under tensor product operations. The inherent noncommutative situation in the case of nonsingular polynomial matrices makes things more difficult, especially if concrete isomorphisms are to be constructed. Of particular difficulty is the lack of a concrete representation of $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2}$. This is the result of the absence of a nice representation of a two-sided greatest common divisor. The tensor products of function spaces given by (5.15), (5.16), and (5.17) just set the stage. In studying tensor products of polynomial or rational models, there are essentially four ways to proceed. For nonsingular polynomial matrices $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[z]^{m\times m}$, one can study the $\mathbb{F}$ – and $\mathbb{F}[z]$ – tensor products, i.e., $X_{D_1} \otimes_{\mathbb{F}} X_{D_2}$ and $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2}$, respectively. Additionally, one can study the polynomial, and rational, models defined by the Kronecker products $D_1(z) \otimes D_2^\top(w)$ and $D_1(z) \otimes D_2^\top(z)$, respectively. Later on, it will be shown that the $\mathbb{F}$-tensor product $X_{D_1} \otimes_{\mathbb{F}} X_{D_2}$ and the polynomial model $X_{D_1(z)\otimes D_2^\top(w)}$ are isomorphic, which will reduce the complexity to the study of three distinct spaces.

Before starting the analysis of the four different tensor product representations, some useful notation and terminology must be established. For rectangular matrices $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{k\times \ell}$, and in this chapter only, the **Kronecker product** is defined as the $mk \times n\ell$ matrix

$$A \otimes B = \begin{pmatrix} b_{11}A & \dots & b_{1\ell}A \\ \vdots & \ddots & \vdots \\ b_{k1}A & \dots & b_{k\ell}A \end{pmatrix}.$$

Note that this definition is in harmony with the definition of the tensor product $f \otimes g$ of two linear maps $f : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ and $g : \mathbb{F}^\ell \longrightarrow \mathbb{F}^k$. In fact, if A and B denote the matrices of f and g with respect to the standard basis, then $A \otimes B$ is the matrix representation of $f \otimes g$ with respect to the standard basis. Nevertheless, we warn the reader that this definition of the Kronecker product is slightly different from that used by many other authors, in the sense that what we denote as $A \otimes B$ is usually denoted as $B \otimes A$. We will use the preceding definition of the Kronecker product only in this chapter, in order to simplify some of the expressions. Later on, in Part III of this book, we will return to the standard definition of the Kronecker product.

By the definition of the Kronecker product, the Kronecker product of an upper triangular matrix A with an rectangular matrix B is block-upper triangular. In particular, the Kronecker product $B \otimes I_N$ is of the form

$$I_N \otimes B = \begin{pmatrix} b_{11}I_N & \dots & b_{1\ell}I_N \\ \vdots & \ddots & \vdots \\ b_{k1}I_N & \dots & b_{k\ell}I_N \end{pmatrix},$$

while

$$A \otimes I_N = \operatorname{diag}(A, \dots, A) = \begin{pmatrix} A & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A \end{pmatrix}.$$

For invertible, $n \times n$ and $m \times m$, respectively, matrices A and B, the Kronecker product $A \otimes B$ is invertible, and

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

The following rules for the Kronecker product are easily verified:

$$(A \otimes B) \otimes C = A \otimes (B \otimes C),$$
$$(A \otimes B)(C \otimes D) = AC \otimes BD,$$
$$(A \otimes B)^\top = A^\top \otimes B^\top.$$

Let $\mathrm{vec}(A) \in \mathbb{F}^{mn}$ denote the column vector that is obtained by stacking the second column of A under the first, then the third under the second, and so on. The identity

$$\mathrm{vec}(ABC) = (A \otimes C^\top)\,\mathrm{vec}(B)$$

is very useful in replacing linear matrix equations in vectorized form. Moreover, there exist permutation matrices P and Q such that

$$P(A \otimes B)Q = B \otimes A$$

is true of all matrices A, B. The eigenvalues of $A \otimes B$ are the products $\lambda_i(A)\lambda_j(B)$ of the eigenvalues $\lambda_i(A)$ and $\lambda_j(B)$ of A and B, respectively. Therefore, the trace and determinant of $A \otimes B$ of matrices A and B are $\mathrm{tr}\,(A \otimes B) = \mathrm{tr}\,(A)\mathrm{tr}\,(B)$ and $\det(A \otimes B) = \det(A)^m \det(B)^n$. Similarly, the eigenvalues of $A \otimes I_m + I_n \otimes B$ are the sums $\lambda_i(A) + \lambda_j(B)$.

1. Kronecker Product Polynomial Models

Our aim is to obtain concrete representations of the tensor products $X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top}$ and $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top}$. To this end, the theory of polynomial and rational models is extended to the case of models induced by Kronecker products of polynomial matrices in one or two variables. Polynomial models based on Kronecker product representations were first studied by Helmke and Fuhrmann (1998) in order to obtain explicit descriptions for tangent spaces for manifolds of rational transfer functions.

Recalling the identification (5.17) and the fact that $\mathbb{F}((z^{-1}))^p$ is a vector space over the field $\mathbb{F}((z^{-1}))$ allows us to introduce a module structure on the space of truncated matrix Laurent series in two variables, z and w, i.e., on $\mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$, as follows.

Definition 5.12. For Laurent series $A_1(z) \in \mathbb{F}((z^{-1}))^{p\times p}$ and $A_2(w) \in \mathbb{F}((w^{-1}))^{m\times m}$, define their $\mathbb{F}$**-Kronecker product** $A_1(z) \otimes A_2^\top(w)$ as the map

$$(A_1(z) \otimes A_2^\top(w)) : \mathbb{F}((z^{-1}, w^{-1}))^{p\times m} \longrightarrow \mathbb{F}((z^{-1}, w^{-1}))^{p\times m}$$
$$(A_1(z) \otimes A_2^\top(w))F(z,w) = A_1(z)F(z,w)A_2(w).$$

Clearly, $A_1(z) \otimes A_2^\top(w)$ is an $\mathbb{F}((z^{-1}, w^{-1}))$-linear map and, hence, also an $\mathbb{F}$-linear map. Similarly, one defines the $\mathbb{F}[z]$**-Kronecker product** $A_1(z) \otimes A_2^\top(z)$ as the map

$$(A_1(z) \otimes A_2^\top(z)) : \mathbb{F}((z^{-1}))^{p\times m} \longrightarrow \mathbb{F}((z^{-1}))^{p\times m}$$
$$(A_1(z) \otimes A_2^\top(z))F(z) = A_1(z)F(z)A_2(z).$$

There are many derivatives of this definition. In particular, we will look at the restriction to polynomial spaces $\mathbb{F}[z]^{p\times m}$ and $\mathbb{F}[z,w]^{p\times m}$, i.e., to spaces of polynomial matrices in one or two variables. Thus, we define the two projection maps $\pi_{D_1(z)\otimes D_2^\top(w)} : \mathbb{F}[z,w]^{p\times m} \longrightarrow \mathbb{F}[z,w]^{p\times m}$ and $\pi_{D_1(z)\otimes D_2^\top(z)} : \mathbb{F}[z]^{p\times m} \longrightarrow \mathbb{F}[z]^{p\times m}$ by

$$\begin{aligned}\pi_{D_1(z)\otimes D_2^\top(w)}F(z,w) &= (D_1(z)\otimes D_2^\top(w))(\pi_-^z\otimes\pi_-^w)(D_1(z)\otimes D_2^\top(w))^{-1}F(z,w)\\ &= (\pi_{D_1(z)}\otimes_{\mathbb{F}}\pi_{D_2^\top(w)})F(z,w) \qquad (5.20)\end{aligned}$$

and

$$\begin{aligned}\pi_{D_1(z)\otimes D_2^\top(z)}F(z) &= (D_1(z)\otimes D_2^\top(z))\pi_-(D_1(z)\otimes D_2^\top(z))^{-1}F(z)\\ &= D_1(z)[\pi_-(D_1(z)^{-1}F(z)D_2(z)^{-1})]D_2(z) \qquad (5.21)\\ &= (\pi_{D_1(z)}\otimes_{\mathbb{F}[z]}\pi_{D_2^\top(z)})F(z),\end{aligned}$$

respectively. Clearly, $\pi_-^z\otimes\pi_-^w$ is a projection map in $\mathbb{F}((z^{-1},w^{-1}))^{p\times m}$ and π_- a projection map in $\mathbb{F}((z^{-1}))^{p\times m}$. Hence, $\pi_{D_1(z)\otimes D_2^\top(w)}$ is a projection map in $\mathbb{F}[z,w]^{p\times m}$ and $\pi_{D_1(z)\otimes D_2^\top(z)}$ a projection map in $\mathbb{F}[z]^{p\times m}$. There are two important special cases of these maps, namely,

$$\begin{aligned}\pi_{D_1(z)\otimes_{\mathbb{F}}I}Q(z,w) &= \pi_{D_1(z)}Q(z,w),\\ \pi_{I\otimes_{\mathbb{F}}D_2^\top(w)}Q(z,w) &= Q(z,w)\pi_{D_2^\top(w)}.\end{aligned}$$

To formulate the basic properties of the projection operators, we first prove an elementary result about projections.

Lemma 5.13. *Let $\mathscr{X}$ be a linear space and P_1 and P_2 two commuting linear projections acting in $\mathscr{X}$, i.e., $P_1P_2=P_2P_1$. Then:*

$$\begin{aligned}\mathrm{Ker}P_1P_2 &= \mathrm{Ker}P_1+\mathrm{Ker}P_2,\\ \mathrm{Im}P_1P_2 &= \mathrm{Im}P_1\cap\mathrm{Im}P_2.\end{aligned} \qquad (5.22)$$

Proof. Since $\mathrm{Ker}\,P_1,\mathrm{Ker}\,P_2\subset\mathrm{Ker}\,P_1P_2=\mathrm{Ker}\,P_2P_1$, also $\mathrm{Ker}\,P_1+\mathrm{Ker}\,P_2\subset\mathrm{Ker}\,P_1P_2$. Conversely, assume $x\in\mathrm{Ker}\,P_1P_2$. This implies $P_2x\in\mathrm{Ker}\,P_1$. The representation $x=(x-P_2x)+P_2x$, with $(x-P_2x)\in\mathrm{Ker}\,P_2$ and $P_2x\in\mathrm{Ker}\,P_1$, shows that $\mathrm{Ker}\,P_1P_2\subset\mathrm{Ker}\,P_1+\mathrm{Ker}\,P_2$, and (5.22) follows. By the commutativity assumption, $\mathrm{Im}\,P_1P_2\subset\mathrm{Im}\,P_1\cap\mathrm{Im}\,P_2$. Conversely, assuming $x\in\mathrm{Im}\,P_1\cap\mathrm{Im}\,P_2$, there exist vectors $z,w\in\mathscr{X}$ for which $x=P_1z=P_2w$. Since $P_2^2=P_2$, this implies $x=P_2w=P_2P_1z\in\mathrm{Im}\,P_1P_2$, i.e., $\mathrm{Im}\,P_1\cap\mathrm{Im}\,P_2\subset\mathrm{Im}\,P_1P_2$. The two inclusions imply the equality. ∎

Proposition 5.14. *Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(w)\in\mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices. Then:*

1. *The maps $\pi_{D_1(z)\otimes I}$, $\pi_{I\otimes D_2^\top(w)}$, and $\pi_{D_1(z)}\otimes_{\mathbb{F}}\pi_{D_2^\top(w)}$ are all projections in $\mathbb{F}[z,w]^{p\times m}$;*
2. *The projections $\pi_{D_1(z)\otimes I}$ and $\pi_{I\otimes D_2^\top(w)}$ commute, and*

$$\pi_{D_1(z)\otimes I}\pi_{I\otimes D_2^\top(w)}=\pi_{I\otimes D_2^\top(w)}\pi_{D_1(z)\otimes I}=\pi_{D_1(z)\otimes D_2^\top(w)}=\pi_{D_1(z)}\otimes_{\mathbb{F}}\pi_{D_2^\top(w)}; \qquad (5.23)$$

3. *The following characterizations are valid:*

$$\begin{aligned}
\operatorname{Ker}\pi_{D_1(z)\otimes I} &= D_1(z)\mathbb{F}[z,w]^{p\times m},\\
\operatorname{Ker}\pi_{I\otimes D_2^\top(w)} &= \mathbb{F}[z,w]^{p\times m}D_2(w),\\
\operatorname{Ker}\pi_{D_1(z)\otimes D_2^\top(w)} &= D_1(z)\mathbb{F}[z,w]^{p\times m}+\mathbb{F}[z,w]^{p\times m}D_2(w).
\end{aligned} \tag{5.24}$$

Proof. 1. Follows from the fact that π_{D_1} is a projection in $\mathbb{F}[z]^p$ and π_{D_2} a projection in $\mathbb{F}[z]^m$.

2. From the isomorphism (5.15) it follows that elements of the form $f(z)\otimes_{\mathbb{F}} g(w) = f(z)g^\top(w)$ span $\mathbb{F}[z,w]^{p\times m}$. On elements of this form

$$\begin{aligned}
\pi_{D_1(z)\otimes I}\pi_{I\otimes D_2^\top(w)} f\otimes_{\mathbb{F}} g &= \pi_{D_1(z)\otimes I}(f\otimes_{\mathbb{F}}\pi_{D_2^\top(w)}g)\\
&= (\pi_{D_1(z)}f\otimes_{\mathbb{F}}\pi_{D_2(w)}g) = \pi_{I\otimes D_2^\top(w)}(\pi_{D_1(z)}f\otimes_{\mathbb{F}} g)\\
&= \pi_{I\otimes D_2^\top(w)}\pi_{D_1(z)\otimes I}(f\otimes_{\mathbb{F}} g),
\end{aligned}$$

from which (5.23) follows.

3. Clearly, $Q(z,w)\in\operatorname{Ker}\pi_{D_1(z)\otimes I}$ if and only if $\pi_-^z D_1(z)^{-1}Q(z,w)=0$, i.e., $D_1(z)^{-1}Q(z,w)=P(z,w)$ for some polynomial matrix $P(z,w)$. This is equivalent to $Q(z,w)=D_1(z)P(z,w)\in D_1(z)\mathbb{F}[z,w]^{p\times m}$. The second equality is proved analogously. The third equality follows from Lemma 5.13 and the commutativity of the projections $\pi_{D_1(z)\otimes I}$ and $\pi_{I\otimes D_2^\top(w)}$. ∎

Definition 5.15. Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(w)\in\mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices.

1. The **two-variable Kronecker product polynomial model** is defined by

$$X_{D_1(z)\otimes D_2^\top(w)} = \operatorname{Im}\pi_{D_1(z)\otimes D_2^\top(w)} \subset \mathbb{F}[z,w]^{p\times m}, \tag{5.25}$$

   where the projection $\pi_{D_1(z)\otimes D_2^\top(w)}$ is defined by (5.20).

2. The **Kronecker product polynomial model** is defined by

$$X_{D_1(z)\otimes D_2^\top(z)} = \operatorname{Im}\pi_{D_1(z)\otimes D_2^\top(z)} \subset \mathbb{F}[z]^{p\times m}, \tag{5.26}$$

   where the projection $\pi_{D_1(z)\otimes D_2^\top(z)}$ is defined by (5.21).

Note that in either of these cases, the spaces $X_{D_1(z)\otimes D_2^\top(w)}$ and $X_{D_1(z)\otimes D_2^\top(z)}$ can be identified with polynomial models for the Kronecker products $D_1(z)\otimes D_2^\top(w)$ and $D_1(z)\otimes D_2^\top(z)$, respectively.

Theorem 5.16. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(w) \in \mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices. Then:*

1. *$Q(z,w) \in \mathbb{F}[z,w]^{p\times m}$ satisfies $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ if and only if the rational matrix function*

$$D_1(z)^{-1}Q(z,w)D_2(w)^{-1} \in z^{-1}\mathbb{F}[z^{-1},w^{-1}]^{p\times m}w^{-1}$$

is strictly proper in both variables;

2. *The set $J = D_1(z)\mathbb{F}[z,w]^{p\times m} + \mathbb{F}[z,w]^{p\times m}D_2(w)$ is an $\mathbb{F}[z,w]$-submodule of $\mathbb{F}[z,w]^{p\times m}$. The following isomorphism of $\mathbb{F}[z,w]$-torsion modules is valid:*

$$\begin{aligned} X_{D_1(z)\otimes D_2^\top(w)} &\simeq \mathbb{F}[z,w]^{p\times m}/(D_1(z)\mathbb{F}[z,w]^{p\times m} + \mathbb{F}[z,w]^{p\times m}D_2(w)) \\ &= X_{D_1(z)\otimes I} \cap X_{I\otimes D_2^\top(w)} \\ &\simeq X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top}; \end{aligned} \tag{5.27}$$

3. *The following dimension formula is valid:*

$$\dim X_{D_1(z)\otimes D_2^\top(w)} = \deg(\det D_1)\cdot\deg(\det D_2);$$

4. *A polynomial matrix satisfies $Q(z) \in X_{D_1(z)\otimes D_2^\top(z)}$ if and only if $D_1(z)^{-1}Q(z)$ $D_2(z)^{-1}$ is strictly proper;*
5. *One has*

$$\operatorname{Ker}\pi_{D_1(z)\otimes D_2^\top(z)} = D_1(z)\mathbb{F}[z]^{p\times m}D_2(z), \tag{5.28}$$

and $D_1(z)\mathbb{F}[z]^{p\times m}D_2(z)$ is a full submodule of $\mathbb{F}[z]^{p\times m}$. Hence, there is an isomorphism

$$X_{D_1(z)\otimes D_2^\top(z)} \simeq \mathbb{F}[z]^{p\times m}/(D_1(z)\mathbb{F}[z]^{p\times m}D_2(z)),$$

with both sides being $\mathbb{F}[z]$-torsion modules;

6. *The following dimension formula is valid:*

$$\dim X_{D_1(z)\otimes D_2^\top(z)} = \deg(\det D_1)\cdot\deg(\det D_2).$$

Proof. 1. A $p\times m$ polynomial matrix $Q(z,w)$ is in $X_{D_1(z)\otimes D_2^\top(w)}$ if and only if $Q(z,w) = \pi_{D_1(z)\otimes D_2^\top(w)}Q(z,w)$. In view of (5.20), this is equivalent to

$$(D_1(z)\otimes D_2^\top(w))^{-1}Q(z,w) = (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z)\otimes D_2^\top(w))^{-1}Q(z,w),$$

i.e., to $D_1(z)^{-1}Q(z,w)D_2(w)^{-1} \in z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$.

2. By Proposition 5.14, the map

$$\pi_{D_1(z)} \otimes_{\mathbb{F}} \pi_{D_2^\top(w)} = \pi_{D_1(z)\otimes D_2^\top(w)} : \mathbb{F}[z,w]^{p\times m} \longrightarrow X_{D_1(z)\otimes D_2^\top(w)}$$

is surjective and $\mathbb{F}[z,w]$-linear, with $\operatorname{Ker}\pi_{D_1(z)\otimes D_2^\top(w)} = J$. Thus, the first isomorphism in (5.27) holds. The second equality follows from Lemma 5.13, while the third isomorphism follows directly from Proposition 5.8.

Clearly, J is an $\mathbb{F}[z,w]$-submodule of $\mathbb{F}[z,w]^{p\times m}$. Using, with $d(z) = \det D(z)$, the identity $d(z)I = D(z)\operatorname{adj} D(z)$, we get the inclusion $d_1(z)\mathbb{F}[z,w]^{p\times m}d_2(w) \subset D_1(z)\mathbb{F}[z,w]^{p\times m} + \mathbb{F}[z,w]^{p\times m}D_2(w)$. In turn, this implies that

$$\pi_{D_1(z)\otimes D_2^\top(w)}(d_1(z)d_2(w)Q(z,w)) = 0$$

for all $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$, i.e., the quotient module $\mathbb{F}[z,w]^{p\times m}/J$ is an $\mathbb{F}[z,w]$-torsion module.

3. Follows from (5.27), using the fact that the dimension of the $\mathbb{F}$-tensor product of two $\mathbb{F}$-vector spaces is the product of their dimensions.
4. Follows trivially from (5.21).
5. Clearly, $D_1(z)\mathbb{F}[z]^{p\times m}D_2(z) \subset \operatorname{Ker}\pi_{D_1(z)\otimes D_2^\top(z)}$. Conversely, assume that $Q(z) \in \operatorname{Ker}\pi_{D_1(z)\otimes D_2^\top(z)}$. By the invertibility of the multiplication operator $D_1(z)\otimes D_2^\top(z)$, this means $\pi_- D_1(z)^{-1}Q(z)D_2(z)^{-1} = 0$. Thus there exists a $P(z) \in \mathbb{F}[z]^{p\times m}$ such that $D_1(z)^{-1}Q(z)D_2(z)^{-1} = P(z)$ or $Q(z) = D_1(z)P(z)D_2(z)$, which implies the inclusion $\operatorname{Ker}\pi_{D_1\otimes_{\mathbb{F}[z]}D_2^\top} \subset D_1(z)\mathbb{F}[z]^{p\times m}D_2(z)$, and hence (5.28) follows.
6. Without loss of generality, we can assume that $D_1(z)$ and $D_2(z)$ are in Smith canonical form with invariant factors $d_1^{(1)}(z),\dots,d_p^{(1)}(z)$ and $d_1^{(2)}(z),\dots,d_m^{(2)}(z)$, respectively. Thus, the quotient module $\mathbb{F}[z]^{p\times m}/(D_1(z)\mathbb{F}[z]^{p\times m}D_2(z))$ is a finite-dimensional $\mathbb{F}$-vector space of dimension

$$\sum_{i=1}^{p}\sum_{j=1}^{m} \deg d_i^{(1)} \deg d_j^{(2)} = \deg\det D_1(z) \deg\det D_2(z).$$

This completes the proof. ∎

2. Tensored Rational Models

In analogy with the introduction of tensored polynomial models, we introduce the tensored rational models. Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(w) \in \mathbb{F}[w]^{m\times m}$ be nonsingular.

Define a projection map

$$\pi^{D_1(z)\otimes D_2^\top(w)} : z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1} \longrightarrow z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$$

by

$$\pi^{D_1(z)\otimes D_2^\top(w)}H(z,w)$$
$$= (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z)\otimes D_2^\top(w))^{-1}(\pi_+^z \otimes_{\mathbb{F}} \pi_+^w)(D_1(z)\otimes D_2^\top(w))H(z,w).$$

The **two-variable Kronecker product rational model** and the **Kronecker product rational model** are defined as

$$\begin{aligned} X^{D_1(z)\otimes D_2^\top(w)} &:= \operatorname{Im}\pi^{D_1(z)\otimes D_2^\top(w)} \subset z^{-1}\mathbb{F}[z^{-1},w^{-1}]^{p\times m}w^{-1}, \\ X^{D_1(z)\otimes D_2^\top(z)} &:= \operatorname{Im}\pi^{D_1(z)\otimes D_2^\top(z)} \subset z^{-1}\mathbb{F}[z^{-1}]^{p\times m}, \end{aligned} \tag{5.29}$$

respectively.

Equation (5.29) provides an image representation of the rational model $X^{D_1(z)\otimes D_2^\top(w)}$. To derive a kernel representation of rational models, we introduce **two-variable Toeplitz operators** on $z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$. For $P_1(z)\in\mathbb{F}[z]^{p\times p}$ and $P_2(w)\in\mathbb{F}[w]^{m\times m}$, we define the **Toeplitz operator** $P_1(\sigma)\otimes_{\mathbb{F}} P_2^\top(\tau)$, acting on a truncated Laurent series in two variables $H(z,w)\in z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$, by

$$(P_1(\sigma)\otimes_{\mathbb{F}} P_2^\top(\tau))H(z,w) = (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(P_1(z)H(z,w)P_2(w)).$$

Special cases are the backward shifts σ and τ in the variables z and w, respectively.

Lemma 5.17. *Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(w)\in\mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices. Let $H(z,w)\in z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]^{p\times m}w^{-1}$. A necessary and sufficient condition for $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$ to be valid is $D_1(z)H(z,w)D_2(w)\in\mathbb{F}[z,w]^{p\times m}$. This is equivalent to the kernel representation*

$$X^{D_1(z)\otimes D_2^\top(w)} = \operatorname{Ker}(D_1(\sigma)\otimes_{\mathbb{F}} D_2^\top(\tau)). \tag{5.30}$$

Proof. Clearly, $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$ if and only if $H(z,w)=\pi^{D_1(z)\otimes D_2^\top(w)}H(z,w)$. Assume $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$. Computing

$$\begin{aligned} &(\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z)\otimes D_2^\top(w))H(z,w) \\ &= (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z)\otimes D_2^\top(w))\pi^{D_1(z)\otimes D_2(w)}H(z,w) \\ &= (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(\pi_+^z \otimes_{\mathbb{F}} \pi_+^w)(D_1(z)\otimes D_2^\top(w))H(z,w) = 0 \end{aligned}$$

implies $X^{D_1(z)\otimes D_2^\top(w)} \subset \operatorname{Ker}(D_1(\sigma)\otimes_{\mathbb{F}} D_2^\top(\tau))$.

Conversely, assuming $H(z,w)\in\operatorname{Ker}(D_1(\sigma)\otimes_{\mathbb{F}} D_2^\top(\tau))$ implies $D_1(z)H(z,w)$ $D_2(w)\in\mathbb{F}[z,w]^{p\times m}$. This implies

$$(\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z) \otimes D_2^\top(w))^{-1}(\pi_+^z \otimes_{\mathbb{F}} \pi_+^w)(D_1(z) \otimes D_2^\top(w))H(z,w)$$
$$= (\pi_-^z \otimes_{\mathbb{F}} \pi_-^w)(D_1(z) \otimes D_2^\top(w))^{-1}(D_1(z) \otimes D_2^\top(w))H(z,w) = H(z,w),$$

i.e., $\mathrm{Ker}\,(D_1(\sigma) \otimes_{\mathbb{F}} D_2^\top(\tau)) \subset X^{D_1(z)\otimes D_2^\top(w)}$. The two inclusions imply (5.30). ∎

The elements of $X^{D_1(z)\otimes D_2^\top(w)}$ are rational functions of a special type. They are characterized next.

Proposition 5.18. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(w) \in \mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices. Every element $H(z,w) \in X^{D_1(z)\otimes D_2^\top(w)}$ is a rational, strictly proper function in two variables that has a representation of the form*

$$H(z,w) = \frac{P(z,w)}{d_1(z)d_2(w)}, \tag{5.31}$$

with $d_i(z) = \det D_i(z) \in \mathbb{F}[z]$ nonzero polynomials and $P(z,w) \in \mathbb{F}[z,w]^{p\times m}$.

Proof. From Lemma 5.17 it follows that $H(z,w) \in X^{D_1(z)\otimes D_2^\top(w)}$ if and only if $Q(z,w) = D_1(z)H(z,w)D_2(w) \in \mathbb{F}[z,w]^{p\times m}$. Letting $d_1(z) = \det D_1(z)$ and $d_2(w) = \det D_1(w)$, we compute

$$H(z,w) = D_1(z)^{-1}Q(z,w)D_2(w)^{-1} = \frac{\mathrm{adj}\, D_1(z)Q(z,w)\mathrm{adj}\, D_2(w)}{d_1(z)d_2(w)} = \frac{P(z,w)}{d_1(z)d_2(w)}.$$ ∎

Rational functions of the form (5.31) are called **separable** and have the property that the set of poles is a direct product $A \times B$ of two finite subsets of the algebraic closure $\overline{\mathbb{F}}$. We refer to Fliess (1970) for a characterization of rational elements of $\mathbb{F}[[z^{-1}, w^{-1}]]$ in terms of a finite rank condition of an appropriate Hankel matrix.

The study of duality for the tensor product of models is our next topic.

Theorem 5.19. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[z]^{m\times m}$ be nonsingular. Then*

$$(X_{D_1(z)\otimes D_2^\top(z)})^* \simeq (\mathbb{F}[z]^{p\times m}/D_1(z)\mathbb{F}[z]^{p\times m}D_2(z))^* \simeq X^{D_1^\top(z)\otimes D_2(z)}$$

and

$$(X_{D_1(z)\otimes D_2^\top(w)})^* \simeq X^{D_1^\top(z)\otimes D_2(w)} \simeq X_{D_1^\top(z)\otimes D_2(w)}$$

are $\mathbb{F}[z]$- and $\mathbb{F}[z,w]$-linear isomorphisms, respectively.

Proof. Proposition 5.16 implies the following isomorphism of $\mathbb{F}[z]$-modules:

$$(X_{D_1(z)\otimes D_2^\top(z)})^* \simeq (\mathbb{F}[z]^{p\times m}/D_1(z)\mathbb{F}[z]^{p\times m}D_2(z))^*.$$

Using the vec-operator, we can identify the matrix space $\mathbb{F}[z]^{p\times m}$ with the space of vector polynomials $\mathbb{F}[z]^{pm}$. Moreover, the identity

$$(D_1(z) \otimes D_2^\top(z))^\top = D_1^\top(z) \otimes D_2(z)$$

is satisfied. For vectors of polynomials $F(z)$ and strictly proper functions $H(z)$ let $[F,H] = (F^\top(z)H(z))_{-1}$, denote the residue term of $F(z)^\top H(z)$. Using Theorem 3.38, the map

$$X^{D_1^\top(z)\otimes_{\mathbb{F}[z]}D_2(z)} \longrightarrow (X_{D_1(z)\otimes D_2^\top(z)})^*$$
$$H(z) \mapsto \Big(F(z) \mapsto \mathrm{Trace}(F(z)^\top H(z))_{-1}\Big)$$

yields the module isomorphism

$$(X_{D_1(z)\otimes D_2^\top(z)})^* \simeq X^{D_1^\top(z)\otimes_{\mathbb{F}[z]}D_2(z)}.$$

This completes the proof for the first isomorphisms.

By Proposition 5.16, the elements of the rational model $X^{D_1^\top(z)\otimes D_2(w)}$ are of the form $H(z,w) = D_1(z)^{-\top}Q(z,w)D_2(w)^{-\top}$. Consider the map

$$L : X^{D_1^\top(z)\otimes D_2(w)} \longrightarrow (X_{D_1(z)\otimes D_2^\top(w)})^* \tag{5.32}$$

defined by

$$L(H)F := [F(z,w),H(z,w)] = \mathrm{Trace}(F(z,w)^\top H(z,w))_{-1,-1}.$$

Here, $[\cdot,\cdot]$ denotes the dual pairing on functions of two variables, defined by (5.18). To prove that this pairing is nondegenerate, assume that a polynomial $F(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ satisfies $[F,H] = 0$ for all $H \in X^{D_1^\top(z)\otimes D_2(w)}$. Equivalently,

$$[D_1(z)^{-1}F(z,w)D_2(w)^{-1}, Q(z,w)] = 0$$

for all polynomial matrices Q of the appropriate size. By (5.19), this is equivalent to

$$D_1(z)^{-1}F(z,w)D_2(w)^{-1} \in z^{-1}\mathbb{F}[[z^{-1},w^{-1}]]w^{-1} \cap \left(\mathbb{F}[z,w^{-1}]]^{p\times m} + \mathbb{F}[[z^{-1},w]^{p\times m}\right)$$
$$= \{0\}.$$

Thus, $F(z,w) = 0$. Similarly, $[F,H] = 0$ for all $F \in X_{D_1(z)\otimes D_2^\top(w)}$ implies $H = 0$. Thus the bilinear form $[\cdot,\cdot]$ on $X_{D_1(z)\otimes D_2^\top(w)} \times X^{D_1^\top(z)\otimes D_2(w)}$ is nondegenerate and therefore induces a vector space isomorphism (5.32). Moreover, the natural $\mathbb{F}[z,w]$-module action on, respectively, $X_{D_1(z)\otimes D_2^\top(w)}$ and $X^{D_1^\top(z)\otimes D_2(w)})$ implies

$$[p(z,w)\cdot F,H] = [p(z,w)F,H] = [F,p(z,w)H] = [F,p(z,w)\cdot H].$$

Therefore, (5.32) defines an $\mathbb{F}[z,w]$-linear isomorphism. The equality

$$X^{D_1^\top(z)\otimes D_2(w)} = D_1(z)^{-\top}\otimes D_2(w)^{-1}X_{D_1^\top(z)\otimes D_2(w)}$$

implies the $\mathbb{F}[z,w]$-linear isomorphism

$$X^{D_1^\top(z)\otimes D_2(w)} \simeq X_{D_1^\top(z)\otimes D_2(w)}.$$

This completes the proof. ∎

4. $\mathbb{F}$-Tensored Polynomial Models

Our attention turns to the study of $\mathbb{F}$-tensor products of vectorial polynomial models. Proposition 5.16 implies that a Kronecker tensored polynomial model, in the sense of (5.25), is isomorphic to the tensor product of polynomial models taken over the field $\mathbb{F}$. This is no longer true if tensored polynomial models in the sense of (5.26) are used, and indeed, the models $X_{D_1(z)\otimes D_2^\top(z)}$ and $X_{D_1(z)}\otimes_{\mathbb{F}} X_{D_2(z)}$ are generally not isomorphic. The next proposition gives a concrete, functional representation of the $\mathbb{F}$-tensor product of two polynomial models.

Proposition 5.20. *Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(z)\in\mathbb{F}[z]^{m\times m}$ be nonsingular polynomial matrices. Let $\phi : X_{D_1}\times X_{D_2^\top}\longrightarrow X_{D_1}\otimes_{\mathbb{F}} X_{D_2^\top}$ be the canonical isomorphism, and let $\gamma : X_{D_1}\times X_{D_2^\top}\longrightarrow X_{D_1(z)\otimes D_2^\top(w)}$ be $\mathbb{F}$-bilinear.*

1. *The map $\gamma_* : X_{D_1}\otimes_{\mathbb{F}} X_{D_2^\top}\longrightarrow X_{D_1(z)\otimes D_2^\top(w)}$ defined by*

$$\gamma_*(f_1\otimes_{\mathbb{F}} f_2) = f_1(z)f_2^\top(w)$$

is an $\mathbb{F}[z,w]$-linear isomorphism implying

$$X_{D_1}\otimes_{\mathbb{F}} X_{D_2^\top} \simeq X_{D_1(z)\otimes D_2^\top(w)}. \tag{5.33}$$

In particular, this gives a concrete representation of the tensor product.

2. *The following dimension formula is valid:*

$$\dim(X_{D_1(z)}\otimes_{\mathbb{F}} X_{D_2^\top(w)}) = \deg(\det D_1)\cdot\deg(\det D_2).$$

Proof. 1. Noting the isomorphism (5.27), we compute

$$\begin{aligned} X_{D_1(z)\otimes D_2^\top(w)} &\simeq \mathbb{F}[z,w]^{p\times m}/(D_1(z)\mathbb{F}[z,w]^{p\times m}+\mathbb{F}[z,w]^{p\times m}D_2(w)) \\ &\simeq (\mathbb{F}[z]^p/D_1(z)\mathbb{F}[z]^p)\otimes_{\mathbb{F}}(\mathbb{F}[z]^m/D_2^\top(z)\mathbb{F}[z]^m) \\ &\simeq X_{D_1(z)}\otimes_{\mathbb{F}} X_{D_2^\top(z)}. \end{aligned}$$

2. Using the dimension formula $\dim X_D = \deg(\det D)$ and the isomorphism (5.33), we obtain for the $\mathbb{F}$-tensor product of two polynomial models the dimension formula

$$\dim\left(\mathbb{F}[z]^p/D_1(z)\mathbb{F}[z]\otimes_{\mathbb{F}}\mathbb{F}[w]^m/D_2^\top(w)\mathbb{F}[w]^m\right) = \deg(\det D_1)\cdot\deg(\det D_2). \qquad \blacksquare$$

Note that the polynomial models X_{D_1} and X_{D_2} not only have a vector space structure but are actually $\mathbb{F}[z]$-modules. This implies that $X_{D_1(z)\otimes D_2^\top(w)}$ and, hence, using the isomorphism (5.33), $X_{D_1}\otimes_{\mathbb{F}} X_{D_2^\top}$ have natural $\mathbb{F}[z,w]$-module structures. This is defined by

$$p(z,w)\cdot Q(z,w) = \pi_{D_1(z)\otimes D_2^\top(w)}p(z,w)Q(z,w), \qquad Q(z,w)\in X_{D_1(z)\otimes D_2^\top(w)}, \tag{5.34}$$

where $p(z,w)\in\mathbb{F}[z,w]$.

Similarly, we define an $\mathbb{F}[z,w]$-module structure on the tensored rational model $X^{D_1(z)\otimes D_2^\top(w)}$ by letting,

$$p(z,w)\cdot H(z,w) = \pi^{D_1(z)\otimes D_2^\top(w)}[\textstyle\sum_{i=1}^k\sum_{j=1}^l p_{ij}z^{i-1}H(z,w)w^{j-1}] \tag{5.35}$$

for $p(z,w)=\sum_{i=1}^k\sum_{j=1}^l p_{ij}z^{i-1}w^{j-1}\in\mathbb{F}[z,w]$ and $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$.

Proposition 5.21. *Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(w)\in\mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices and $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$.*

1. *The $\mathbb{F}[z,w]$-module structure on $X^{D_1(z)\otimes D_2^\top(w)}$ defined by (5.35) can be rewritten as*

$$p(z,w)\cdot H(z,w) = (\pi_-^z\otimes\pi_-^w)\sum_{i=1}^k\sum_{j=1}^l p_{ij}z^{i-1}H(z,w)w^{j-1}. \tag{5.36}$$

2. *With the $\mathbb{F}[z,w]$-module structure on $X_{D_1(z)\otimes D_2^\top(w)}$ and $X^{D_1(z)\otimes D_2^\top(w)}$, given by (5.36) and (5.35) respectively, the multiplication map*

$$D_1(z)\otimes D_2^\top(w) : X^{D_1(z)\otimes D_2^\top(w)}\longrightarrow X_{D_1(z)\otimes_{\mathbb{F}} D_2^\top(w)}$$

is an $\mathbb{F}[z,w]$-module isomorphism, giving

$$X_{D_1(z)\otimes D_2^\top(w)}\simeq X^{D_1(z)\otimes D_2^\top(w)}.$$

Proof. 1. Follows from (5.35).

2. Follows, using Lemma 5.17, from the fact that $H(z,w)\in X^{D_1(z)\otimes D_2^\top(w)}$ if and only if $H(z,w)=\pi^{D_1(z)\otimes D_2^\top(w)}H(z,w)$.

Equivalently, if and only if $\pi_{D_1(z)\otimes D_2^\top(w)} D_1(z)H(z,w)D_2(w) = D_1(z)H(z,w)$ $D_2(w)$, i.e., $D_1(z)H(z,w)D_2(w) \in X_{D_1(z)\otimes D_2^\top(w)}$. ∎

Special cases of interest are the single-variable shift operators

$$S_z, S_w : X_{D_1(z)\otimes D_2^\top(w)} \longrightarrow X_{D_1(z)\otimes D_2^\top(w)},$$

defined by

$$\begin{aligned} S_z Q(z,w) &= \pi_{D_1(z)\otimes D_2^\top(w)} zQ(z,w) = \pi_{D_1(z)} zQ(z,w), \\ S_w Q(z,w) &= \pi_{D_1(z)\otimes D_2^\top(w)} Q(z,w)w = \pi_{I\otimes_{\mathbb{F}} D_2^\top(w)} Q(z,w)w. \end{aligned}$$

A concrete representation of the dual space to a tensored polynomial model is given next. For subspaces $\mathscr{U}$ and $\mathscr{V}$ of a linear space $\mathscr{X}$, we shall use the isomorphism $(\mathscr{X}/\mathscr{V})^* \simeq \mathscr{V}^\perp$, as well as the identity $(\mathscr{U}+\mathscr{V})^\perp = \mathscr{U}^\perp \cap \mathscr{V}^\perp$.

Theorem 5.22. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(w) \in \mathbb{F}[w]^{m\times m}$ be nonsingular polynomial matrices. Then the following is an $\mathbb{F}[z,w]$-linear isomorphism:*

$$(X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top})^* \simeq X^{D_1(z)\otimes_{\mathbb{F}} I} \cap X^{I\otimes_{\mathbb{F}} D_2^\top(w)} \simeq X^{D_1^\top(z)\otimes D_2(w)}.$$

Proof. By Proposition 5.16 and Theorem 5.19, the following are $\mathbb{F}[z,w]$-linear isomorphisms:

$$\begin{aligned} X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top} &\simeq X_{D_1(z)\otimes D_2^\top(w)}, \\ (X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top})^* &\simeq X^{D_1^\top(z)\otimes D_2(w)}. \end{aligned}$$

This implies the $\mathbb{F}[z,w]$-linear isomorphisms

$$\begin{aligned} (X_{D_1} \otimes_{\mathbb{F}} X_{D_2^\top})^* &\simeq (\mathbb{F}[z,w]^{m\times p}/(D_1(z)\mathbb{F}[z,w]^{p\times m} + \mathbb{F}[z,w]^{p\times m}D_2(w)))^* \\ &\simeq (D_1(z)\mathbb{F}[z,w]^{p\times m} + \mathbb{F}[z,w]^{p\times m}D_2(w))^\perp \\ &= (D_1(z)\mathbb{F}[z,w]^{p\times m})^\perp \cap (\mathbb{F}[z,w]^{p\times m}D_2(w))^\perp \\ &= X^{D_1(z)\otimes_{\mathbb{F}} I} \cap X^{I\otimes_{\mathbb{F}} D_2^\top(w)}. \end{aligned}$$

Here, the identities

$$\left(D_1^\top(z)\mathbb{F}[z,w]^{p\times m}\right)^\perp = X^{D_1(z)\otimes_{\mathbb{F}} I},$$
$$\left(\mathbb{F}[z,w]^{p\times m}D_2^\top(w)\right)^\perp = X^{I\otimes_{\mathbb{F}} D_2^\top(w)}$$

were used. They follow from the duality relation based on the bilinear form (5.18). Indeed, $H(z,w)\in(D_1(z)\mathbb{F}[z,w]^{p\times m})^\perp$ if and only if, for every $Q(z,w)\in\mathbb{F}[z,w]^{p\times m}$,

$$0=[D_1(z)Q(z,w),H(z,w)]=[Q(z,w),D_1(z)H(z,w)],$$

i.e., if and only if $D_1(z)H(z,w)\in\mathbb{F}[z,w]^{p\times m}$, which implies $H(z,w)\in X^{D_1(z)\otimes I}$. The other formula is proved similarly. ■

3. $\mathbb{F}[z]$-tensored polynomial models

In the preceding parts, tensor product representations of polynomial models over a field were studied. Things change dramatically when tensor products of polynomial models are taken over the polynomial ring $\mathbb{F}[z]$. This leads directly to the study of intertwining maps, the Sylvester equation, and, in a very natural way, to the study of generalized Bezoutians.

Definition 5.23. Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(z)\in\mathbb{F}[z]^{m\times m}$ be nonsingular polynomial matrices. The **greatest common left Kronecker divisor** of $D_1(z)$ and $D_2(z)$ is defined as the greatest common left divisor $D_1(z)\wedge D_2(z)$ of the polynomial matrices $D_1(z)\otimes I_m$ and $I_p\otimes D_2(z)^\top$.

Of course, by construction, the greatest common left Kronecker divisor $D_1(z)\wedge D_2(z)\in\mathbb{F}[z]^{pm\times pm}$ is a nonsingular polynomial matrix. Further elementary properties of $D_1(z)\wedge D_2(z)$ are listed subsequently in Corollary 5.25. The $\mathbb{F}[z]$-tensor product of the polynomial models X_{D_1} and X_{D_2} is characterized by the following theorem. It shows in particular that the $\mathbb{F}[z]$-tensor product of two polynomial models is isomorphic to a polynomial model, defined by the greatest common left Kronecker divisor.

Theorem 5.24. *Let $D_1(z)\in\mathbb{F}[z]^{p\times p}$ and $D_2(z)\in\mathbb{F}[z]^{m\times m}$ be nonsingular polynomial matrices. Let J be the submodule of $\mathbb{F}[z]^{p\times m}$ defined by*

$$J=D_1(z)\mathbb{F}[z]^{p\times m}+\mathbb{F}[z]^{p\times m}D_2(z).$$

1. The $\mathbb{F}[z]$-tensor product can be identified by the following isomorphism:

$$X_{D_1}\otimes_{\mathbb{F}[z]}X_{D_2^\top}\simeq\mathbb{F}[z]^{p\times m}/(D_1(z)\mathbb{F}[z]^{p\times m}+\mathbb{F}[z]^{p\times m}D_2(z)).$$

The isomorphism is given by the canonical map $\gamma : X_{D_1} \times X_{D_2} \longrightarrow X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2}$ *defined by*

$$\gamma(f_1, f_2) = [f_1 f_2^\top]_J,$$

where $[F]_J$ *denotes the equivalence class of* $F \in \mathbb{F}[z]^{p\times m}$ *with respect to submodule* J.

2. *Let* $D_1 \wedge D_2 \in \mathbb{F}[z]^{pm\times pm}$ *denote the greatest common left divisor of* $D_1(z) \otimes I_m$ *and* $I_p \otimes D_2^\top(z)$. *Then* $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top}$ *is* $\mathbb{F}[z]$*-linearly isomorphic to the polynomial model*

$$X_{D_1 \wedge D_2}.$$

Proof. The first claim follows trivially from Proposition 5.3.
By identifying $\mathbb{F}[z]^{p\times m}$ with $\mathbb{F}[z]^{pm}$ we obtain the module isomorphism

$$X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top} \simeq \mathbb{F}[z]^{pm}/D(z)\mathbb{F}[z]^{2pm},$$

where $D(z) := (D_1(z) \otimes I_m, I_p \otimes D_2^\top(z)) \in \mathbb{F}[z]^{pm\times 2pm}$. Thus

$$D(z) = (D_1(z) \wedge D_2(z))A(z),$$

where $A(z) \in \mathbb{F}[z]^{pm\times 2pm}$ is left prime. Thus $A(z)\mathbb{F}[z]^{2pm} = \mathbb{F}[z]^{pm}$, and therefore

$$D(z)\mathbb{F}[z]^{2pm} = D_1(z) \wedge D_2(z)\mathbb{F}[z]^{pm}.$$

This implies $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top} \simeq \mathbb{F}[z]^{pm}/D_1(z) \wedge D_2(z)\mathbb{F}[z]^{pm} \simeq X_{D_1 \wedge D_2}$. This completes the proof of the second part. ■

Corollary 5.25. *Let* $D_1(z) \in \mathbb{F}[z]^{p\times p}$ *and* $D_2(z) \in \mathbb{F}[z]^{m\times m}$ *be nonsingular polynomial matrices with invariant factors* $d_1^{(1)}(z), \ldots, d_p^{(1)}(z)$ *and* $d_1^{(2)}(z), \ldots, d_m^{(2)}(z)$, *respectively. Let* $d_i^{(1)} \wedge d_j^{(2)}$ *denote the greatest common divisor of the polynomials* $d_i^{(1)}(z)$ *and* $d_j^{(2)}(z)$.

1. *The tensor product* $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2}$ *is isomorphic to the polynomial model* $X_{D_1 \wedge D_2}$.
2. *For scalar polynomials* $d_1(z), d_2(z)$ *there is an isomorphism of* $\mathbb{F}[z]$*-modules*

$$X_{d_1} \otimes_{\mathbb{F}[z]} X_{d_2} \simeq X_{d_1 \wedge d_2}.$$

3. $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2}$ *is* $\mathbb{F}[z]$*-linearly isomorphic to the direct sum*

$$\bigoplus_{i=1}^{p} \bigoplus_{j=1}^{m} X_{d_i^{(1)} \wedge d_j^{(2)}}.$$

In particular, the following dimension formula is valid:

$$\dim X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2} = \sum_{i=1}^{p}\sum_{j=1}^{m} \deg(d_i^{(1)} \wedge d_j^{(2)}).$$

4. *The invariant factors of* $D_1(z) \wedge D_2(z)$ *are* $d_i^{(1)}(z) \wedge d_j^{(2)}(z)$. *In particular,*

$$\det D_1(z) \wedge D_2(z) = \det D_1(z)^m \det D_2(z)^p.$$

Proof. Assertions 1 and 2 follow at once from part 2 of Theorem 5.24.

3. Let $\Delta_1 = \mathrm{diag}\,(d_1^{(1)},\ldots,d_p^{(1)})$ and $\Delta_2 = \mathrm{diag}\,(d_1^{(2)},\ldots,d_m^{(2)})$ be the respective Smith forms of $D_1(z)$ and $D_2(z)$, and let $U_i(z)$ and $V_i(z)$ be unimodular polynomial matrices satisfying $U_i(z)D_i(z) = \Delta_i(z)V_i(z)$. This implies the $\mathbb{F}[z]$-linear isomorphism

$$X_{\Delta_1} \simeq \bigoplus_{i=1}^{p} X_{d_i^{(1)}}, \quad X_{\Delta_2} \simeq \bigoplus_{j=1}^{m} X_{d_j^{(2)}}.$$

Using 2, the isomorphisms

$$\begin{aligned} X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2} &\simeq X_{\Delta_1} \otimes_{\mathbb{F}[z]} X_{\Delta_2} \simeq \left(\bigoplus_{i=1}^{p} X_{d_i^{(1)}}\right) \otimes_{\mathbb{F}[z]} \left(\bigoplus_{i=1}^{m} X_{d_j^{(2)}}\right) \\ &\simeq \bigoplus_{i,j} X_{d_i^{(1)}} \otimes_{\mathbb{F}[z]} X_{d_j^{(2)}} \simeq \bigoplus_{i=1}^{p}\bigoplus_{j=1}^{m} X_{d_i^{(1)} \wedge d_j^{(2)}} \end{aligned}$$

follow. This completes the proof of 3.

4. Consider the unimodular polynomial matrices $U_i(z)$ and $V_i(z)$ such that

$$D_1(z) = U_1(z)\Delta_1(z)V_1(z), \quad D_2(z)^\top = U_2(z)\Delta_2(z)^\top V_2(z)$$

are in Smith form. Since $D_1 \wedge D_2$ is the greatest common left factor of $D_1 \otimes I_m$ and $I_p \otimes D_2^\top$ there exist polynomial matrices $M(z)$ and $N(z)$ such that

$$\begin{aligned} (D_1 \wedge D_2)M &= D_1 \otimes I_m = (U_1 \otimes U_2)(\Delta_1 \otimes I)(V_1 \otimes U_2^{-1}), \\ (D_1 \wedge D_2)N &= (U_1 \otimes U_2)(I_p \otimes \Delta_2^\top)(U_1^{-1} \otimes V_2). \end{aligned}$$

Thus there exist unimodular matrices $P(z) = (U_1 \otimes U_2)^{-1}$, $R(z) = (V_1^{-1} \otimes U_2)$, and $S(z) = (U_1 \otimes V_2^{-1})$ such that

$$P(D_1 \wedge D_2)MR = \Delta_1 \otimes I, \quad P(D_1 \wedge D_2)NS = I_p \otimes \Delta_2^\top.$$

This implies that $P(z)(D_1 \wedge D_2)$ is a greatest common left divisor of $\Delta_1 \otimes I$ and $I_p \otimes \Delta_2^\top$ and therefore, up to an irrelevant unimodular factor, coincides with $\Delta_1 \wedge \Delta_2$. Thus $D_1 \wedge D_2$ and $\Delta_1 \wedge \Delta_2$ have the same invariant factors. It is easy to see that $\Delta_1 \wedge \Delta_2$ can be chosen as a diagonal matrix with diagonal entries $d_i^{(1)} \wedge d_j^{(2)}$. This completes the proof. ■

Corollary 5.26. *Consider nonsingular polynomial matrices $D(z) \in \mathbb{F}[z]^{p\times p}$ and $\overline{D}(z) \in \mathbb{F}[z]^{m\times m}$ with the same nontrivial invariant factors d_i, ordered so that $d_i | d_{i-1}$. Then the following assertions hold:*

1.

$$\dim \mathrm{Hom}_{\mathbb{F}[z]}(S_{\overline{D}}, S_D) = \sum_i (2i-1)\deg d_i.$$

2. Let $A \in \mathbb{F}^{n\times n}$ have invariant factors $d_1,\dots,d_n$ ordered such that $d_i|d_{i-1}$. Let $\mathscr{C}(A) = \{X \in \mathbb{F}^{n\times n} \mid AX = XA\}$ denote the centralizer of A. Then:

$$\dim \mathscr{C}(A) = \sum_{i=1}^{n} (2i-1)\deg d_i; \tag{5.37}$$

3. For $A \in \mathbb{F}^{n\times n}$ with invariant factors $d_1,\dots,d_n$, $\dim \mathscr{C}(A) = n^2$ if and only if there exists an $\alpha \in \mathbb{F}$ such that, for all $i = 1,\dots,n$, $d_i(z) = z-\alpha$. Equivalently, $A = \alpha I$, i.e., A is a scalar transformation.

The relation of the $\mathbb{F}[z]$-tensor product $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top}$ to the tensored Kronecker model $X_{D_1(z)\otimes D_2^\top(z)}$ is examined next.

Proposition 5.27. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[z]^{m\times m}$ be nonsingular. The following is an $\mathbb{F}[z]$-linear isomorphism:*

$$X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top} \simeq X_{D_1(z)\otimes D_2^\top(z)} / (D_1 X_{I\otimes D_2^\top} + X_{D_1\otimes I} D_2). \tag{5.38}$$

Proof. With i the canonical injections, the following diagram is commutative. Here $\pi_{D_1\otimes D_2^\top}|J$ denotes the restriction of $\pi_{D_1\otimes D_2^\top}$ to the subspace $J = D_1(z)\mathbb{F}[z]^{p\times m} + \mathbb{F}[z]^{p\times m}D_2(z)$:

$$\begin{array}{ccc}
D_1(z)\mathbb{F}[z]^{p\times m} + \mathbb{F}[z]^{p\times m}D_2(z) & \xrightarrow{\quad i \quad} & \mathbb{F}[z]^{p\times m} \\
\Big\downarrow {\scriptstyle \pi_{D_1\otimes D_2^\top}|J} & & \Big\downarrow {\scriptstyle \pi_{D_1\otimes D_2^\top}} \\
D_1 X_{I\otimes D_2^\top} + X_{D_1\otimes I} D_2 & \xrightarrow{\quad i \quad} & X_{D_1\otimes D_2^\top}
\end{array}$$

Moreover, $\pi^{-1}_{D_1\otimes D_2^\top}(D_1 X_{I\otimes D_2^\top} + X_{D_1\otimes I} D_2) = D_1(z)\mathbb{F}[z]^{p\times m} + \mathbb{F}[z]^{p\times m} D_2(z)$. Using the surjectivity of $\pi_{D_1\otimes D_2^\top} : \mathbb{F}[z]^{p\times m} \longrightarrow X_{D_1\otimes D_2^\top}$ and applying a standard argument, the isomorphism (5.38) follows. ∎

The next theorem yields an explicit description of the module $(X_{D_1} \otimes_{\mathbb{F}[z]} X^*_{D_2})^*$.

Theorem 5.28. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[z]^{m\times m}$ be nonsingular. The following are $\mathbb{F}[z]$-linear isomorphisms:*

$$X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top} \simeq (X_{D_1} \otimes_{\mathbb{F}[z]} X^*_{D_2})^* \simeq (X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top})^* \simeq X^{D_1\otimes I} \cap X^{I\otimes D_2^\top}.$$

Proof. For each submodule $M \subset \mathbb{F}[z]^{p\times m}$, one can identify the annihilator with $M^\perp = \{H \in z^{-1}\mathbb{F}[[z^{-1}]]^{p\times m} \mid \mathrm{Trace}[F,H] = 0 \ \forall F \in M\}$. Here $[F,H] = (F(z)^\top H(z))_{-1}$ denotes the residue. Computing

$$\begin{aligned}(X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top})^* &\simeq \Big(\mathbb{F}[z]^{p\times m}/(D_1(z)\mathbb{F}[z]^{p\times m} + \mathbb{F}[z]^{p\times m} D_2^\top(z))\Big)^* \\ &= (D_1(z)\mathbb{F}[z]^{p\times m} + \mathbb{F}[z]^{p\times m} D_2^\top(z))^\perp \\ &= (D_1(z)\mathbb{F}[z]^{p\times m})^\perp \cap (\mathbb{F}[z]^{p\times m} D_2^\top(z))^\perp \\ &= X^{D_1\otimes I} \cap X^{I\otimes D_2^\top}\end{aligned}$$

proves one isomorphism.

To prove the other isomorphisms, one uses the Smith form. Thus, for each nonsingular polynomial matrix $Q(z)$, the invariant factors of $Q(z)$ and $Q^\top(z)$ are equal, implying the $\mathbb{F}[z]$-linear isomorphism

$$X_Q \simeq X_{Q^\top}.$$

Moreover, Theorem 3.38 implies the isomorphism $X_Q \simeq X_Q^*$. By Theorem 5.24, the tensor product $X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top}$ is $\mathbb{F}[z]$-linearly isomorphic to the polynomial model $X_{D_1\wedge D_2^\top}$ and hence also to $X^*_{D_1\wedge D_2^\top} \simeq (X_{D_1} \otimes_{\mathbb{F}[z]} X_{D_2^\top})^*$. This completes the proof. ∎

5.3 Polynomial Sylvester Equation

Proceeding now to a more detailed study of the Sylvester equation in the tensored polynomial model framework, definition (5.34) is specialized to the polynomial $p(z,w) = z - w$. One obtains, for all $Q(z,w) \in X_{D_1(z)\otimes_{\mathbb{F}} D_2^\top(w)}$, that

$$\mathscr{S} Q(z,w) = (z-w)\cdot Q(z,w) = \pi_{D_1(z)\otimes D_2^\top(w)}(zQ(z,w) - Q(z,w)w). \qquad (5.39)$$

The map $\mathscr{S}$ is referred to as the **polynomial Sylvester operator** . In fact, with constant matrices $A_1 \in \mathbb{F}^{p\times p}$ and $A_2 \in \mathbb{F}^{m\times m}$, and defining $D_1(z) = zI - A_1$ and $D_2(w) = wI - A_2$, we obtain $X_{D_1(z)\otimes_{\mathbb{F}} D_2^\top(w)} = \mathbb{F}^{p\times m}$. Therefore, $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ if and only if $Q(z,w) \in \mathbb{F}^{p\times m}$, i.e., $X = Q(z,w)$ is a constant matrix. This implies

$$(z-w)\cdot X = \pi_{(zI-A_1)\otimes(wI-A_2^\top)}(z-w)X = A_1X - XA_2$$

for all $X \in \mathbb{F}^{p\times m}$, i.e., we recover the standard **Sylvester operator**. This computation shows that the **classical Sylvester equation**

$$A_1X - XA_2 = C$$

corresponds to the equation

$$\mathscr{S}X = C,$$

with $X, C \in X_{(zI-A_1)\otimes_{\mathbb{F}}(wI-A_2^\top)}$ necessarily constant matrices.

Note that every polynomial matrix $T(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ has a factorization of the form

$$T(z,w) = R_1(z)R_2^\top(w),$$

with $R_1(z) \in X_{D_1\otimes I_k} \subset \mathbb{F}[z]^{p\times k}$ and $R_2(w)^\top \in X_{I_k\otimes D_2^\top(w)} \subset \mathbb{F}[w]^{k\times m}$, and both polynomial matrices $R_1(z)$ and $R_2(w)$ have linearly independent columns. The following theorem reduces the analysis of the general Sylvester equation to a polynomial equation of the Bezout type. This extends the method, introduced in Willems and Fuhrmann (1992), for the analysis of the Lyapunov equation. Of course, a special case is the homogeneous Sylvester equation, which has a direct connection to the theory of Bezoutians.

Theorem 5.29. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(w) \in \mathbb{F}[w]^{m\times m}$ be nonsingular, let the Sylvester operator $\mathscr{S} : X_{D_1(z)\otimes D_2^\top(w)} \longrightarrow X_{D_1(z)\otimes D_2^\top(w)}$ be defined by (5.39) and let $R_1(z) \in X_{D_1(z)\otimes_{\mathbb{F}} I}$, $R_2^\top(w) \in X_{I\otimes_{\mathbb{F}} D_2^\top(w)}$. Then:*

1. The Sylvester equation

$$S_{D_1}Q - QS_{D_2} = T(z,w) = R_1(z)R_2^\top(w), \tag{5.40}$$

or equivalently

$$\mathscr{S}Q(z,w) = \pi_{D_1(z)\otimes D_2^\top(w)}(z-w)Q(z,w) = R_1(z)R_2^\top(w),$$

is solvable if and only if there exists polynomial matrices $N_1(z) \in X_{D_1(z)\otimes I}$ *and* $N_2(z) \in X_{I\otimes D_2^\top(z)}$ *for which*

$$D_1(z)N_2(z) - N_1(z)D_2(z) + R_1(z)R_2^\top(z) = 0. \tag{5.41}$$

Equation (5.40) will be referred to as the **polynomial Sylvester equation***, or* **PSE** *for short, and its solutions are given by*

$$Q(z,w) = \frac{D_1(z)N_2(w) - N_1(z)D_2(w) + R_1(z)R_2^\top(w)}{z-w}; \tag{5.42}$$

2. $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ *solves the* **homogeneous polynomial Sylvester equation***, or* **HPSE** *for short, if and only if there exist polynomial matrices* $N_1(z) \in X_{D_1\otimes I}$ *and* $N_2(z) \in X_{I\otimes D_2^\top}$ *that satisfy*

$$D_1(z)N_2(z) - N_1(z)D_2(z) = 0, \tag{5.43}$$

in terms of which

$$Q(z,w) = \frac{D_1(z)N_2(w) - N_1(z)D_2(w)}{z-w}.$$

Proof. 1. Assume there exist the polynomial matrices $N_1(z) \in X_{D_1\otimes I}$ and $N_2(z) \in X_{I\otimes D_2^\top}$, solving equation (5.41), and for which $Q(z,w)$ is defined by (5.42). Note first that, under our assumptions on $R_1(z)$ and $R_2(w)$,

$$\begin{aligned} &D_1(z)^{-1}Q(z,w)D_2(w)^{-1} \\ &= \frac{N_2(w)D_2(w)^{-1} - D_1(z)^{-1}N_1(z) + D_1(z)^{-1}R_1(z)R_2^\top(w)D_2(w)^{-1}}{z-w} \end{aligned}$$

is strictly proper in both variables, i.e., $Q(z,w)$ is in $X_{D_1(z)\otimes D_2^\top(w)}$. Computing

$$\begin{aligned} \mathscr{S}Q(z,w) &= \pi_{D_1(z)\otimes D_2^\top(w)}(z-w)Q(z,w) \\ &= \pi_{D_1(z)\otimes D_2^\top(w)}(D_1(z)N_2(w) - N_1(z)D_2(w) + R_1(z)R_2(w)) \\ &= R_1(z)R_2(w)^\top, \end{aligned}$$

it follows that $Q(z,w)$ is indeed a solution.

To prove the converse, note that, given a nonsingular polynomial matrix $D_1(z) \in \mathbb{F}[z]^{p\times p}$, then, for $f(z) \in X_{D_2}$, $(S_{D_2}f)(z) = zf(z) - D_2(z)\xi_f$, where $\xi_f = (D_2^{-1}f)_{-1}$. This implies that, for $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$,

$$S_{z\otimes 1}Q(z,w) = zQ(z,w) - D_1(z)N_2(w), \quad S_{1\otimes w}Q(z,w) = Q(z,w)w - N_1(z)D_2(w),$$

with both $N_2D_2^{-1}$ and $D_1^{-1}N_1$ strictly proper. Assuming $Q(z,w)$ is a solution of the PSE, we compute

$$\begin{aligned} S_{z-w}Q(z,w) &= [zQ(z,w) - D_1(z)N_2(w)] - [Q(z,w)w - N_1(z)D_2(w)] \\ &= R_1(z)R_2(w)^\top, \end{aligned}$$

implying

$$Q(z,w) = \frac{D_1(z)N_2(w) - N_1(z)D_2(w) + R_1(z)R_2(w)^\top}{z-w}.$$

However, because $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ is a polynomial matrix, (5.41) necessarily holds.

2. Follows from the previous part. ■

This leads us to introduce the following object.

Definition 5.30. A polynomial matrix $Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ is called a **generalized Bezoutian** if it has a representation of the form

$$Q(z,w) = \frac{D_1(z)N_2(w) - N_1(z)D_2(w)}{z-w}, \tag{5.44}$$

with $D_1^{-1}N_1$ and $N_2D_2^{-1}$ strictly proper and such that the identity

$$D_1(z)N_2(z) = N_1(z)D_2(z) \tag{5.45}$$

is satisfied.

Corollary 5.31. *$Q(z,w) \in X_{D_1(z)\otimes D_2^\top(w)}$ is a solution of the HPSE (5.43) if and only if $Q(z,w)$ is a generalized Bezoutian.*

Proof. Follows from Theorem 5.29.2. ■

5.4 Generalized Bezoutians and Intertwining Maps

Proposition 5.6 shows that there is a close connection between tensor products of vector spaces and spaces of $\mathbb{F}$-linear maps between vector spaces. For functional models one can be more specific about the form of such connections, leading to a new interpretation of Bezoutian operators in terms of intertwining maps and module homomorphisms of polynomial models. Denote by $\mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2})$ the space of all $\mathbb{F}$-linear maps from X_{D_1} to X_{D_2} and by $\mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})$ the space of all $\mathbb{F}[z]$-

linear maps from X_{D_1} to X_{D_2}, i.e., the space of all $\mathbb{F}$-linear maps Z from X_{D_1} to X_{D_2} that satisfy $ZS_{D_1} = S_{D_2}Z$. The essential information that encodes the mutual interrelations between these spaces is given by the following commutative diagram:

$$\begin{array}{ccc} X_{D_2} \otimes_{\mathbb{F}} X_{D_1^\top} & \xrightarrow{\quad \Psi \quad} & \mathrm{Hom}\,_{\mathbb{F}}(X_{D_1}, X_{D_2}) \\ \uparrow \beta & & \uparrow i \\ X_{D_2} \otimes_{\mathbb{F}[z]} X_{D_1^\top} & \xrightarrow{\quad \psi \quad} & \mathrm{Hom}\,_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}) \end{array}$$

Here i is the natural inclusion of $\mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})$ in $\mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2})$. The map β will be constructed via Bezoutians, and we will establish the $\mathbb{F}$- and $\mathbb{F}[z]$-linear isomorphisms Ψ and ψ, respectively. As is easily seen, this diagram is an equivalent reformulation of the first diagram, mentioned at the beginning of this chapter, insofar as the dual spaces $X_{D_1}^*$ and $(X_{D_2} \otimes X_{D_1}^*)^*$ are replaced by their isomorphic counterparts $X_{D_1^\top}$ and $X_{D_2} \otimes X_{D_1^\top}$, respectively. For the relevant isomorphisms that underpin such reformulations, we refer to Theorem 5.28. It may come as somewhat of a surprise that the study of $\mathbb{F}[z]$-homomorphisms of polynomial models can be based on the study of tensored models. This (see Theorem 5.34) leads to a further clarification of the connection between intertwining maps and Bezoutians.

To achieve an even more concrete form of the previous diagram, we next prove a proposition that establishes a concrete connection between the space of maps Z intertwining the shifts S_{D_2} and S_{D_1} and the $\mathbb{F}[z]$-tensor product of the polynomial models $X_{D_2^\top}$ and X_{D_1}.

Proposition 5.32. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[w]^{m\times m}$ be nonsingular.*

1. *Every $Q(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$ has a representation of the form*

$$Q(z,w) = R_2(z)R_1^\top(w), \tag{5.46}$$

with $R_2(z) \in X_{D_2(z)\otimes_{\mathbb{F}} I}$ and $R_1(w)^\top \in X_{I\otimes_{\mathbb{F}} D_1(w)^\top}$, i.e., both $D_2(z)^{-1}R_2(z)$ and $D_1(w)^{-\top}R_1(w)$ are strictly proper. Without loss of generality, assume that the columns of $R_1(z)$ and $R_2(z)$ are linearly independent.

2. *Define a map $\Psi : X_{D_2(z)\otimes_{\mathbb{F}} D_1^\top(w)} \longrightarrow \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2})$, for $f(z) \in X_{D_1}$ and $Q(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$ having the representation (5.46), by*

$$\Psi(Q)f = Z_Q f = < f, Q(z,.)^\top > = [D_1^{-1}f, Q(z,.)^\top] = (Q(z,\cdot)D_1^{-1}f)_{-1}. \tag{5.47}$$

Then Ψ induces the isomorphism

$$X_{D_2(z)\otimes_{\mathbb{F}} D_1^\top(w)} \simeq \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2}).$$

The polynomial $Q(z,w)$ will be called the **representing kernel** *of the map $\Psi(Q)$, and (5.47) will be referred to as its* **kernel representation**.

Proof. 1. Let $\{u_1(z),\dots,u_{n_1}(z)\}$ be a basis for X_{D_2} and $\{v_1(z),\dots,v_{n_2}(z)\}$ be a basis for X_{D_1}. Assume $Q(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$. Since $D_2(z)^{-1}Q(z,w)$ is strictly proper in z, there exist polynomials $g_i(w)$ such that $Q(z,w) = \sum_{i=0}^{n_2} u_i(z)g_i(w)$. Since $Q(z,w)D_1(w)^{-1}$ is strictly proper in the variable w,

$$\sum_{i=0}^{n_2} u_i(z)\pi_+(g_i(w)D_1(w)^{-1}) = 0.$$

In turn, this implies $g_i^\top(w) \in X_{D_1^\top}$ and, hence, the existence of $\alpha_{ij} \in \mathbb{F}$ for which $g_i(w) = \sum_{j=0}^{n_1} \alpha_{ij} v_j(w)^\top$. Thus

$$Q(z,w) = \sum_{i=0}^{n_2}\sum_{j=0}^{n_1} \alpha_{ij} u_i(z) v_j(w)^\top = R^{(2)}(z)AR^{(1)}(w)^\top.$$

Here $R^{(2)} = (u_1(z),\dots,u_{n_1}(z))$, $R^{(1)} = (v_1(z),\dots,v_{n_2}(z))$ and $A = (\alpha_{ij})$. Next, let $r = \mathrm{rank} A$, which implies the existence of a factorization $A = A_2A_1^\top$, with $A_i \in \mathbb{F}^{n_i\times r}$ of full column rank. Redefining the $R^{(i)}(z)$, the statement follows.

2. As elements of the form $u_i(z)(v_j(w))^\top$ generate $X_{D_2(z)\otimes_{\mathbb{F}} D_1^\top(w)}$, we compute for such an element $\Psi(u_i(z)(v_j(w))^\top) = u_i(z) < f, v_j >$. This allows us to compute, for $Q(z,w) = \sum_{i=1}^{k} u_i(z)v_i(w)^\top = R_2(z)R_1^\top(w)$,

$$\begin{aligned}
&\Psi(\sum_{i=1}^{k} u_i(z)v_i(w)^\top)f = \sum_{i=1}^{k} u_i(z) < f, v_i > \\
&= (\sum_{i=1}^{k} u_i(z)v_i(w)^\top D_1(w)^{-1}f(w))_{-1} = (\sum_{i=1}^{k} u_i(z)v_i(w)^\top D_1(w)^{-1}f(w))_{-1} \\
&= < f, \sum_{i=1}^{k} v_i(w)u_i(z)^\top > = < f, Q(z,\cdot)^\top > .
\end{aligned}$$

Clearly, Ψ defined by (5.47) is $\mathbb{F}$-linear. To show the injectivity of Ψ, assume without loss of generality that the columns of $R_1(z)$ are linearly independent and that $\Psi(Q) = 0$, i.e., that for all $g(z) \in X_{D_1}$,

$$0 = < g, Q(z,.)^\top > = R_2(z)(R_1^\top(w)D_1(w)^{-1}g(w))_{-1}.$$

This implies $(R_1^\top(w)D_1(w)^{-1}g(w))_{-1} = 0$ for all $g(z) \in X_{D_1}$. Since the columns of $R_1(z)$ are in $X_{D_1^\top}$, Theorem 3.38 implies that $R_1 = 0$ and, hence, $Q(z,w) = 0$. That Ψ is an isomorphism follows from the equality of dimension. Indeed,

$$\begin{aligned}\dim X_{D_2(z)\otimes_{\mathbb{F}} D_1^\top(w)} &= \dim X_{D_2} \otimes_{\mathbb{F}} X_{D_1^\top} \\ &= \dim X_{D_2} \cdot \dim X_{D_1^\top} = \deg\det D_1 \cdot \deg\det D_2 \\ &= \dim \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2}).\end{aligned}$$

■

The following lemma will be needed in the sequel.

Lemma 5.33. *Let $H(z) \in \mathbb{F}((z^{-1}))^{p\times m}$. Then:*

$$\left(\frac{H(w)}{w-z}\right)_{-1} = \pi_+ H(z).$$

Proof. Let $H(w) = \sum_{k=-\infty}^{n_H} H_{-k} w^k$. One computes

$$\begin{aligned}\left(\frac{H(w)}{w-z}\right)_{-1} &= \left(\sum_{k=-\infty}^{n_H} H_{-k}\frac{w^k}{w-z}\right)_{-1} = \sum_{k=-\infty}^{n_H} H_{-k}\left(\frac{w^k}{w-z}\right)_{-1} \\ &= \sum_{k=-\infty}^{n_H} H_{-k}\sum_{j=0}^{\infty}\left(w^k\frac{z^j}{w^{j+1}}\right)_{-1} = \sum_{k=0}^{n_H} H_{-k} z^k \\ &= \pi_+ H(z).\end{aligned}$$

Here we used

$$\left(w^k\frac{z^j}{w^{j+1}}\right)_{-1} = \begin{cases} 0 & j \neq k \\ z^k & j = k.\end{cases}$$

■

In Theorem 3.20, a characterization of maps intertwining two polynomial models was derived. In fact, already in equation (3.22) there is a clue to the beautiful link between intertwining maps and the theory of generalized Bezoutians. This connection is now formalized in the following theorem, which plays a central role in our analysis. It allows for a second, independent approach to the characterization of homomorphisms between polynomial models and the commutant lifting theorem (Chapter 3).

Theorem 5.34. *Let $D_1(z) \in \mathbb{F}[z]^{p\times p}$ and $D_2(z) \in \mathbb{F}[w]^{m\times m}$ be nonsingular. Let $R_1(z) \in \mathbb{F}[z]^{p\times k}$ and $R_2(z) \in \mathbb{F}[z]^{m\times k}$. Assume that $D_1(z)^{-\top}R_1(z)$ and $D_2(z)^{-1}R_2(z)$ are strictly proper. Then the following statements are equivalent.*

1. *$Q(z,w) = R_2(z)R_1^\top(w)$ is a solution of the HPSE (5.43).*
2. *$Q(z,w) = R_2(z)R_1^\top(w)$ is a Bezoutian, i.e., it has a representation of the form (5.44) and satisfies (5.45).*
3. *The map $Z : X_{D_1} \longrightarrow X_{D_2}$ defined by*

$$\begin{aligned} Zg &= R_2(z) < g, R_1^\top >=< g, Q(z,.)^\top > \\ &= [D_1^{-1}g, (R_2(z)R_1^\top(w))^\top] = (R_2(z)R_1^\top(w)D_1(w)^{-1}g(w))_{-1} \end{aligned} \tag{5.48}$$

satisfies

$$S_{D_2}Z = ZS_{D_1}, \tag{5.49}$$

i.e., it is an intertwining map or, equivalently, an $\mathbb{F}[z]$-homomorphism.

4. *The map $Z : X_{D_1} \longrightarrow X_{D_2}$ has the representation*

$$Zg = \pi_{D_2}N_2 g, \qquad g \in X_{D_1}, \tag{5.50}$$

with $D_1(z), D_2(z)$ satisfying the intertwining relation

$$N_2(z)D_1(z) = D_2(z)N_1(z) \tag{5.51}$$

for some $N_1(z), N_2(z) \in \mathbb{F}[z]^{p\times m}$.

Proof. (1) ⇔ (2)

Follows from Theorem 5.29 and Definition 5.30.

(2) ⇒ (3)

Assume $Q(z,w) = R_2(z)R_1(w)^\top$ is a Bezoutian, i.e., it has a representation of the form (5.44). We compute, for $g(z) \in X_{D_1}$,

$$\begin{aligned} &(S_{D_2}Z - ZS_{D_1})g \\ &= S_{D_2}(R_2(z)R_1(w)^\top D_1(w)^{-1}g(w))_{-1} - R_2(z)(R_1(w)^\top D_1(w)^{-1}S_{D_1}g(w))_{-1} \\ &= \pi_{D_2}(zR_2(z)R_1^\top(w)D_1(w)^{-1}g(w))_{-1} - R_2(z)(R_1(w)^\top D_1(w)^{-1}\pi_{D_1}wg(w))_{-1} \\ &= \pi_{D_2}(zR_2(z)R_1^\top(w)D_1(w)^{-1}g(w))_{-1} - R_2(z)(R_1(w)^\top \pi_- D_1^{-1}wg(w))_{-1} \\ &= \pi_{D_2}(zR_2(z)R_1^\top(w)D_1(w)^{-1}g(w))_{-1} - R_2(z)(R_1(w)^\top D_1(w)^{-1}wg(w))_{-1} \\ &= \pi_{D_2}((D_2(z)N_1(w) - N_2(z)D_1(w))D_1(w)^{-1}g(w))_{-1} \\ &= -\pi_{D_2}(N_2(z))(D_2(w)D_2(w)^{-1}g(w))_{-1}) \\ &= -N_2(z)(g(w))_{-1} = 0, \end{aligned}$$

using the fact that $g(w)$ is a polynomial. This implies (5.49).

(2) ⇒ (4)

Assume first that $Q(z,w)$ is a Bezoutian, i.e., has a representation of the form (5.44). We prove now that Z has the alternative representation (5.50). To this end, using Lemma 5.33, one computes

$$\begin{aligned}
Zg &=< g, Q(z,\cdot)^\top > \\
&= [D_1^{-1} g, Q(z,\cdot)^\top] = \left(Q(z,w)D_1(w)^{-1}g(w)\right)_{-1} \\
&= \left(\frac{D_2(z)N_1(w) - N_2(z)D_1(w)}{z-w} D_1(w)^{-1}g(w)\right)_{-1} \\
&= \left(D_2(z)\frac{N_1(w)D_1(w)^{-1}g(w)}{z-w} - N_2(z)\frac{g(w)}{z-w}\right)_{-1} \\
&= -D_2(z)\pi_+ N_1 D_1^{-1} g + N_2(z)g(z) = N_2(z)g(z) - D_2\pi_+ D_2^{-1} N_2 g \\
&= D_2 \pi_- D_2^{-1} N_2 g = \pi_{D_2} N_2 g.
\end{aligned}$$

(3) ⇒ (2)

Assume that $Z : X_{D_1} \longrightarrow X_{D_2}$, defined by (5.48), is intertwining. For $g(z) \in X_{D_1}$, one computes, using Lemma 5.33, the fact that $\pi_{D_2} R_2 = R_2$, and that a contribution of a polynomial term to the residue $()_{-1}$ is zero,

$$\begin{aligned}
0 &= (S_{D_2} Z - Z S_{D_1})g = S_{D_2} < g, Q(z,.)^\top > - < S_{D_1} g, Q(z,.)^\top > \\
&= S_{D_2}\left(R_2(z)R_1(w)^\top D_1(w)^{-1}g(w)\right)_{-1} - \left(R_2(z)R_1(w)^\top D_1(w)^{-1} S_{D_1} g(w)\right)_{-1} \\
&= \pi_{D_2}\left(zR_2(z)R_1(w)^\top D_1(w)^{-1}g(w)\right)_{-1} - \left(R_2(z)R_1(w)^\top D_1(w)^{-1}(\pi_{D_1} wg(w))\right)_{-1} \\
&= \pi_{D_2}\left(zR_2(z)R_1(w)^\top D_1(w)^{-1}g(w)\right)_{-1} - \left(R_2(z)R_1(w)^\top D_1(w)^{-1} wg(w)\right)_{-1} \\
&= \pi_{D_2}\left(R_2(z)(z-w)R_1(w)^\top D_1(w)^{-1}g(w)\right)_{-1}.
\end{aligned}$$

Since this is true for all $g(z) \in X_{D_1}$, and as it trivially holds for $g(z) \in D_1 \mathbb{F}[z]^m$, it is satisfied for all $g(z) \in \mathbb{F}[z]^m$. Hence, $\pi_{D_2}\left(R_2(z)(z-w)R_1(w)^\top D_1(w)^{-1}\right)$ is a polynomial in both variables. It follows that $(\pi_{D_2} \otimes \pi_{D_1})\left(R_2(z)(z-w)R_1(w)^\top\right) = 0$, i.e., $R_2(z)R_1(w)^\top$ is a solution of the HPSE. Applying Theorem 5.29, it follows that $R_2(z)R_1(w)^\top$ is a Bezoutian.

(4) ⇒ (3)

From representation (5.50) it easily follows that Z is intertwining. Indeed, noting that equality (5.51) implies $N_2 \operatorname{Ker} \pi_{D_2} \subset \operatorname{Ker} \pi_{D_1}$, we compute

$$S_{D_1} Zg - Z S_{D_2} g = \pi_{D_1} z \pi_{D_1} N_2 g - \pi_{D_1} N_2 \pi_{D_2} zg = \pi_{D_1} z N_2 g - \pi_{D_1} N_2 zg = 0.$$

■

Proposition 5.35. *Let $D_1(z) \in \mathbb{F}[z]^{m\times m}$ and $D_2(z) \in \mathbb{F}[z]^{p\times p}$ be nonsingular.*

1. *Every $H \in X^{D_2\otimes I} \cap X^{I\otimes D_1^\top}$ has unique representations*

$$H(z) = D_2(z)^{-1}N_2(z) = N_1(z)D_1(z)^{-1}, \tag{5.52}$$

   with $N_1(z) \in X^{I\otimes D_1^\top}$ and $N_2(z) \in X_{D_2\otimes I}$.

2. *The map $\psi : X^{D_2\otimes I} \cap X^{I\otimes D_1^\top} \longrightarrow \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})$ defined by*

$$\psi(H)g = \pi_{D_2}N_2 g, \qquad g(z) \in X_{D_1} \tag{5.53}$$

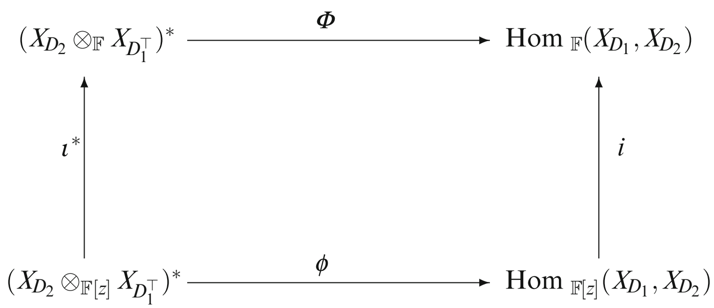
   induces the isomorphism

$$X^{D_2\otimes I} \cap X^{I\otimes D_1^\top} \simeq \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}). \tag{5.54}$$

Proof. 1. That $H(z) \in X^{D_2\otimes I} \cap X^{I\otimes D_1^\top}$ has the unique representations (5.52) is obvious from the definitions.

2. Clearly, by Theorem 5.34, $\psi(H) \in \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})$, i.e., it is an intertwining map. To show the injectivity of the map ψ, assume $\psi(H) = 0$. With the representation (5.52), this implies $\pi_{D_2}N_2 g = 0$ for all $g \in X_{D_1}$. This means that $D_2(z)$ is a left factor of $N_2(z)$. But as $N_2(z) \in X_{D_2\otimes I}$, necessarily $N_2(z) = 0$. That ψ is surjective follows from Theorem 5.34. ∎

The Bezout Map.
We are now in a position to explain the beautiful connection between tensor products, intertwining maps, and Bezoutians in a very concrete way. Our starting point is the commutative diagram of canonical homomorphisms established in the following diagram:



In Proposition (5.6), it was shown that $(X_{D_2} \otimes_{\mathbb{F}} X_{D_1^\top})^*$ is naturally isomorphic to $X_{D_2} \otimes_{\mathbb{F}} X_{D_1^\top}$, which in turn is isomorphic to the Kronecker product space $X_{D_2(z)\otimes D_1^\top(w)}$ of polynomials in two variables z, w. Similarly, using duality theory and Theorem 5.28, the module $(X_{D_2} \otimes_{\mathbb{F}[z]} X_{D_1^\top})^*$ can be identified with $X_{D_2} \otimes_{\mathbb{F}[z]} X_{D_1^\top}$, but now at the expense of a less clear interpretation of the maps ι^* and ϕ.

Matters simplify considerably if, instead of using a polynomial model representation $X_{D_2} \otimes_{\mathbb{F}[z]} X_{D_1^\top}$, we pass to a rational model $X^{D_2 \otimes I} \cap X^{I \otimes D_1^\top}$ via the isomorphism (5.54). This leads to a new commutative diagram of concrete maps that is described now in more detail.

Noting that every $H(z) \in X^{D_2 \otimes I} \cap X^{I \otimes D_1^\top}$ has a unique representation of the form

$$H(z) = D_2(z)^{-1} N_2(z) = N_1(z) D_1(z)^{-1},$$

one defines the **Bezout map**

$$\beta : X^{D_2 \otimes I} \cap X^{I \otimes D_1^\top} \longrightarrow X_{D_2(z) \otimes D_1^\top(w)}$$

by associating with the rational function (z) the two-variable polynomial

$$\beta(H) = Q(z,w) = \frac{D_2(z)N_1(w) - N_2(z)D_1(w)}{z-w}. \tag{5.55}$$

Note that β is $\mathbb{F}$-linear and injective and the image space consists of all **Bezoutian forms**. That $Q(z,w) \in X_{D_2(z) \otimes D_1^\top(w)}$ follows from the calculation

$$\begin{aligned}
&D_2(z)^{-1} Q(z,w) D_1(w)^{-1} = D_2(z)^{-1} \frac{D_2(z)N_1(w) - N_2(z)D_1(w)}{z-w} D_1(w)^{-1} \\
&= \frac{H(w) - H(z)}{z-w} = \sum_{k=1}^{\infty} H_k \frac{w^{-k} - z^{-k}}{z-w} = -\frac{1}{zw} \sum_{k=1}^{\infty} H_k \frac{w^{-k} - z^{-k}}{w^{-1} - z^{-1}} \\
&= -\frac{1}{zw} \sum_{k=1}^{\infty} H_k \left(w^{1-k} + w^{2-k} z^{-1} + \cdots + z^{1-k} \right),
\end{aligned}$$

which shows that it is indeed strictly proper.

With these definitions, the principal result can be stated.

Theorem 5.36. *Let $D_2(z) \in \mathbb{F}[z]^{p \times p}$ and $D_1(z) \in \mathbb{F}[w]^{m \times m}$ be nonsingular. Let the maps*

$$\Psi : X_{D_2(z) \otimes D_1^\top(w)} \longrightarrow \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2})$$

and

$$\psi : X^{D_2 \otimes I} \cap X^{I \otimes D_1^\top} \longrightarrow \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2})$$

be defined by (5.47) and (5.53), respectively. Let

$$i : \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}) \longrightarrow \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2})$$

be the canonical embedding, and let

$$\beta : X^{D_2 \otimes I} \cap X^{I \otimes D_1^\top} \longrightarrow X_{D_2(z) \otimes D_1^\top(w)}$$

be the injective Bezout map given in (5.55). Then the following diagram commutes:

$$\begin{array}{ccc} X_{D_2(z)\otimes D_1^\top(w)} & \xrightarrow{\quad \Psi \quad} & \mathrm{Hom}_{\mathbb{F}}(X_{D_1}, X_{D_2}) \\ \uparrow \beta & & \uparrow i \\ X^{D_2\otimes I} \cap X^{I\otimes D_1^\top} & \xrightarrow{\quad \psi \quad} & \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}) \end{array}$$

Proof. Note that, by Proposition 5.35, every element $H \in X^{D_2\otimes I} \cap X^{I\otimes D_1^\top}$ has a unique representation of the form (5.52). With the **Bezout map** defined by (5.55), all that remains is to prove the identity

$$\Psi \circ \beta = i \circ \psi. \tag{5.56}$$

To this end, with $H(z) \in X^{D_2\otimes I} \cap X^{I\otimes D_1^\top}$, $g \in X_{D_1}$, and using equation (5.50) and Lemma 5.4, we compute

$$\begin{aligned}
(\Psi \circ \beta)(H)g &= \Psi(\beta(H))g = \Psi(Q)g \\
&= < g, Q(z,\cdot)^\top > = \left(\frac{D_2(z)N_1(w) - N_2(z)D_1(w)}{z-w} D_1(w)^{-1} g(w) \right)_{-1} \\
&= D_2(z) \left(\frac{D_2(w)^{-1}N_2(w)g(w)}{z-w} \right)_{-1} - N_2(z) \left(\frac{g(w)}{z-w} \right)_{-1} \\
&= -D_2(z)\pi_+(D_2(z)^{-1}N_2(z)g(z)) + N_2(z)\pi_+(g(z)) \\
&= D_2(z)\pi_-(D_2(z)^{-1}N_2(z)g(z)) - N_2(z)g(z) + N_2(z)\pi_+(g(z)) \\
&= \pi_{D_2} N_2 g = \psi(H)g = (i \circ \psi)(H)g,
\end{aligned}$$

i.e., (5.56) is proved. ■

Theorem 5.36 shows that the $\mathbb{F}$-linear maps $X_{D_1} \longrightarrow X_{D_2}$

$$Zf = \left(Q(z,w) D_1(w)^{-1} f(w) \right)_{-1},$$

whose representing kernel $Q(z,w)$ is a Bezoutian form, correspond exactly to the $\mathbb{F}[z]$-linear homomorphisms $X_{D_1} \longrightarrow X_{D_2}$. Thus the theory of Bezoutians is intimately connected to the structure of module homomorphisms of polynomial models. This is a two-way street because it shows the existence of concrete representations of module homomorphisms by Bezoutian forms. Conversely, the linear maps defined by Bezoutians are seen to be module homomorphisms whose structure is clarified by the commutant lifting theorem, Theorem 2.54.

Generalized Bezoutian Matrices.
Generalized Bezoutian matrices $B(N_1,D_1,N_2,D_2)$, induced by a quadruple of polynomial matrices satisfying an intertwining relation, were introduced in Anderson and Jury (1976) and studied in further detail in Lerer and Tismenetsky (1982). In the sequel we will find it convenient to distinguish between the Bezoutian form as a matrix polynomial in two variables, which is an element of a tensored model, the corresponding intertwining map, and the Bezoutian matrix, which is a specific matrix representation. This is analogous to the distinction between a linear transformation and its matrix representation. There are many choices of bases in polynomial models, and some lead to interesting matrix representations; see, for example, Fuhrmann and Datta (1989) or Mani and Hartwig (1997).

Definition 5.37. Assume that the polynomial matrices $N_1(z)$ and $N_2(z)$ are such that $H(z) = N_1(z)D_1(z)^{-1} = D_2(z)^{-1}N_2(z) \in \mathbb{F}(z)^{p\times m}$ are strictly proper. Let the polynomial matrix $Q(z,w)$ be given as

$$Q(z,w) = \frac{D_2(z)N_1(w) - N_2(z)D_1(w)}{z-w} = \sum_{i,j=1}^{k} Q_{ij} z^{i-1} w^{j-1}.$$

Then the matrix

$$B(N_1,D_1,N_2,D_2) = (Q_{ij}) \ \in \mathbb{F}^{kp\times km}$$

is called the **generalized Bezoutian matrix**. The linear operator

$$\mathbf{B} : X_{D_1} \longrightarrow X_{D_2},\ \mathbf{B}f = \left(Q(z,w)D_1(w)^{-1}f(w)\right)_{-1}$$

is called the **Bezout operator** of $H(z) = N_1(z)D_1(z)^{-1} = D_2(z)^{-1}N_2(z)$.

In Chapter 3, coprimeness conditions for the injectivity, surjectivity, and bijectivity of homomorphisms between polynomial models were obtained. Using Theorem 5.36, this result can now be applied to characterize full rank properties of the Bezout operator.

Theorem 5.38. *The Bezout operator* $\mathbf{B}$ *of* $H(z) = N_1(z)D_1(z)^{-1} = D_2(z)^{-1}N_2(z)$ *is*

1. *Full column rank if and only if* $D_1(z)$ *and* $N_1(z)$ *are right coprime,*
2. *Full row rank if and only if* $D_2(z)$ *and* $N_2(z)$ *are left coprime,*
3. *Invertible if and only if* $D_1(z)$ *and* $N_1(z)$ *are right coprime and* $D_2(z)$ *and* $N_2(z)$ *are left coprime.*

Proof. By Theorem 5.36, the Bezout operator **B** coincides with the homomorphism $Z : X_{D_1} \longrightarrow X_{D_2}$ defined by $Zf = \pi_{D_2}(N_2 f)$. The result follows by applying Theorem 3.21. ■

The Bezout operator has the advantage of providing a simple rank test for coprimeness of pairs of polynomial matrices. However, to compute a matrix representation, basis vectors in the polynomial model spaces must be chosen. In contrast, the generalized Bezoutian matrix by Anderson and Jury (1976) is directly defined as the matrix $(Q_{k\ell})$ of coefficients of the Bezoutian polynomial

$$Q(z,w) = \sum_{k,\ell} Q_{k\ell} z^{k-1} w^{\ell-1}.$$

This matrix is certainly easier to compute than a matrix representation of the Bezout operator, but it is more difficult to explore the structural properties of this matrix. For instance, using Theorem 5.36, the greatest common left and right divisors of polynomial matrices can be characterized in terms of the kernel of the Bezout operator. For the generalized Bezoutian matrix $(Q_{k\ell})$ of Anderson and Jury (1976), this is much harder to achieve. Thus our preference is to work with the foregoing definition.

The strength of the preceding approach is illustrated by briefly discussing the case of classical Bezoutians. Thus, let

$$q(z) = z^n + q_{n-1} z^{n-1} + \cdots + q_0 \in \mathbb{F}[z]$$

be a scalar monic polynomial, and let

$$p(z) = p_{n-1} z^{n-1} + p_{n-2} z^{n-2} + \cdots + p_0 \in \mathbb{F}[z].$$

The **Bezoutian form** then has the expansion

$$\frac{q(z)p(w) - p(z)q(w)}{z - w} = \sum_{i,j=1}^{n} b_{ij} z^{i-1} w^{j-1},$$

with unique coefficients $b_{ij} \in \mathbb{F}$. The **Bezoutian matrix**, then, is the $n \times n$ matrix

$$B(p,q) = (b_{ij}) \in \mathbb{F}^{n \times n}.$$

The following basic representation theorem for Bezoutian matrices is discussed next. Let $\mathscr{B}_{\mathrm{st}} = \{1, z, \ldots, z^{n-1}\}$ denote the **standard basis** of X_q, and let $\mathscr{B}_{\mathrm{co}} = \{e_1(z), \ldots, e_n(z)\}$, with

$$e_i(z) = z^{n-i} + q_{n-1} z^{n-i-1} + \cdots + q_i = \pi_+(z^{-i} q(z)),$$

denote the **control basis** of X_q. One checks that $\mathscr{B}_{\mathrm{co}}$ is the dual basis to $\mathscr{B}_{\mathrm{st}}$ by computing

$$\begin{aligned}\left(\frac{z^{k-1}e_\ell(z)}{q(z)}\right)_{-1} &= \left(\frac{z^{k-1}\pi_+(z^{-\ell}q(z))}{q(z)}\right)_{-1}\\ &= \left(\frac{z^{k-1}(z^{-\ell}q(z))}{q(z)}\right)_{-1} = \left(z^{k-\ell-1}\right)_{-1} = \delta_{k\ell}.\end{aligned}$$

Theorem 5.39. *1. The Bezoutian matrix $B(p,q) = [\mathbf{B}]^{\mathrm{st}}_{\mathrm{co}}$ is the matrix representation of the Bezout operator $\mathbf{B}$ with respect to the control basis and standard basis on X_q.*
2. The Bezoutian can be expressed, using the shift operator $S_q : X_q \longrightarrow X_q$, as

$$B(p,q) = [p(S_q)]^{\mathrm{st}}_{\mathrm{co}}.$$

Proof. Computing

$$\begin{aligned}\mathbf{B}e_j &= \left(\frac{q(z)p(w)-p(z)q(w)}{z-w}q(w)^{-1}e_j(w)\right)_{-1}\\ &= \sum_{r,s=1}^{n} b_{rs}z^{r-1}\left(w^{s-1}q(w)^{-1}\pi_+(w^{-j}q(w))\right)_{-1} = \sum_{r,s=1}^{n} b_{rs}z^{r-1}\left(w^{s-1}q(w)^{-1}(w^{-j}q(w))\right)_{-1}\\ &= \sum_{r,s=1}^{n} b_{rs}z^{r-1}\left(w^{s-j-1}\right)_{-1} = \sum_{r=1}^{n} b_{rj}z^{r-1}\end{aligned}$$

shows that the Bezout matrix $B(p,q) = [\mathbf{B}]^{\mathrm{st}}_{\mathrm{co}}$ is simply a matrix representation of the Bezout operator $\mathbf{B} : X_q \longrightarrow X_q$. This proves the first claim.

Theorem 5.36 implies that, for $j = 1,\dots,n$, $\mathbf{B}e_j = \pi_q(p(z)e_j(z)) = p(S_q)e_j$, which completes the proof. ∎

Recall from (3.26) that the matrix representation of the shift operator S_q with respect to the standard basis on X_q is the companion matrix

$$[S_q]^{\mathrm{st}}_{\mathrm{st}} = C_q := \begin{pmatrix} 0 & 0 & \cdots & & -q_0\\ 1 & 0 & & & -q_1\\ & \ddots & \ddots & & \vdots\\ & & 1 & 0 & -q_{n-2}\\ & & & 1 & -q_{n-1}\end{pmatrix}.$$

Moreover, the basis change matrix $[I]^{\mathrm{st}}_{\mathrm{co}}$ coincides with the Bezoutian $B(1,q)$. Thus the explicit description of the Bezoutian matrix as $B(p,q) = p(C_q)B(1,q)$ is deduced. This expression is often referred to as **Barnett's formula**.

As a further simple consequence, one obtains a classical coprimeness test for scalar polynomials.

Theorem 5.40. *The polynomials $p(z)$ and $q(z)$ are coprime if and only if the Bezoutian matrix $B(p,q)$ is invertible.*

Proof. We know already that the Bezout operator $\mathbf{B} : X_q \longrightarrow X_q$ defined by $h(z) = p(z)q(z)^{-1} = q(z)^{-1}p(z)$ is invertible if and only if p and q are coprime. In the preceding theorem it was shown that $B(p,q)$ is a matrix representation of $\mathbf{B}$. Thus, the result follows. ■

5.5 Stability Characterizations

Characterization of the stability of linear systems is central to systems theory. A discrete-time linear dynamic system

$$x_{t+1} = Ax_t$$

on $\mathbb{F}^n$ is called **asymptotically stable** if the sequences $x_t = A^t x_0$ converge to zero for all initial conditions $x_0 \in \mathbb{F}^n$. Likewise, a continuous-time linear system

$$\dot{x}(t) = Ax(t)$$

is called asymptotically stable whenever

$$\lim_{t\to\infty} e^{tA}x_0 = 0$$

is true for all $x_0 \in \mathbb{F}^n$. Of course, in order for such a notion to make sense, a topology on the field $\mathbb{F}$ must be specified. Throughout this section, we will restrict our discussion to the standard situation where $\mathbb{F}$ denotes either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$, both being endowed with their standard Euclidean topology. To streamline the presentation of the subsequent results, we will mainly restrict ourselves to continuous-time systems and mention corresponding results for discrete-time systems only occasionally.

It is easily seen (see Proposition 5.41 below) that a continuous-time system $\dot{x}(t) = Ax(t)$ is asymptotically stable if and only if all eigenvalues of A have a negative real part. Similarly, a discrete-time system is asymptotically stable if and only if all eigenvalues of A are in the open unit disc. Since the eigenvalues of A are the roots of the characteristic polynomial, determination of the stability of finite-dimensional linear systems reduces to the problem of characterizing the location of zeros of real and complex polynomials. This problem has a long history, and there are two basic approaches to it – via the use of quadratic forms on the one hand or the use of special Sylvester equations on the other. The problem of root location was already solved by Hermite (1856) and Hurwitz (1895) using quadratic form approaches. In this connection one should also refer to the work of Routh (1877) because of the efficiency of the computational algorithm proposed by the researcher. In a major contribution to the subject, Lyapunov (1893) offered a completely different approach based on energy considerations. In the linear case, the Lyapunov theory reduces the study of the stability of a system of first-order homogeneous constant coefficient differential equations to the positive definiteness of the solution of the celebrated Lyapunov equation.

In this section, we will first characterize the asymptotic stability of first-order systems using Lyapunov's method and then discuss the generalization to higher-order systems of differential equations. Because it is trivial to reduce a scalar nth-order homogeneous equation

$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots + a_0 y = 0$$

to a first-order system $\dot{x} = Ax$, it became possible to derive the classical stability criteria for scalar higher-order systems from Lyapunov theory. This was done surprisingly late, and the paper by Parks (1962) is usually considered the first of such derivations. The various reductions seemed to work also for the case of a higher-order system of matrix differential equations of the form

$$y^{(n)} + P_{n-1}y^{(n-1)} + \cdots + P_0 y = 0,$$

i.e., for which the matrix polynomial $P(z) = I_m z^n + P_{n-1}z^{n-1} + \cdots + P_0$ is monic. Strangely, a gap remained in the theory related to finding an algebraic test for the asymptotic stability of solutions of a system of the form

$$P_n y^{(n)} + P_{n-1}y^{(n-1)} + \cdots + P_0 y = 0,$$

where $P(z) = P_n z^n + P_{n-1}z^{n-1} + \cdots + P_0 \in \mathbb{C}^{m\times m}[z]$ is a nonsingular polynomial matrix. It is our aim in this chapter not only to close this gap but to apply the theory of quadratic forms and tensor products of polynomial and rational models to the derivation of stability criteria for higher-order multivariable systems. This leads to the study of two-variable polynomial matrices. It is worth mentioning that Kalman (1969, 1970) utilized a similar idea of switching from a polynomial equation in one variable to a polynomial in two variables and its associated quadratic form. Both these papers deal solely with the scalar case.

1. Lyapunov Stability

A brief summary of the basic facts from linear algebra on complex Hermitian matrices and adjoint operators is presented. In particular, bilinear forms are replaced by **sesquilinear forms**, i.e., forms $< x, y >$ that are antilinear in x and linear in y. Let $\mathscr{X}$ be a finite-dimensional complex vector space, and let $\mathscr{X}^*$ denote its dual space of complex linear functionals on $\mathscr{X}$. In this context there are two different ways of defining duality. The first is the one already employed in Chapter 3, using the canonical dual pairing $\mathscr{X}^* \times \mathscr{X} \longrightarrow \mathbb{C}$ defined by $(\lambda, x) \mapsto \lambda(x)$. The other definition – and this is the one adopted in this section – is given by the **Hermitian**

dual pairing $< \cdot,\cdot >: \mathscr{X}^* \times \mathscr{X} \longrightarrow \mathbb{C}$ defined as

$$< \lambda, x >_{\mathbb{C}} = \overline{\lambda(\bar{x})}.$$

Note that $< \lambda, x >$ is antilinear in λ and linear in x and therefore defines a sequilinear form. Since finite-dimensional vector spaces are reflexive, we can identify $\mathscr{X}^{**}$ with $\mathscr{X}$. Thus

$$< x, \lambda >_{\mathbb{C}} = \overline{< \lambda, x >_{\mathbb{C}}}.$$

Consider now a linear map $T : \mathscr{X} \longrightarrow \mathscr{Y}$, where $\mathscr{X}$ and $\mathscr{Y}$ are complex vector spaces with duals $\mathscr{Y}^*$ and $\mathscr{X}^*$, respectively. The Hermitian adjoint map $T^* : \mathscr{Y}^* \longrightarrow \mathscr{X}^*$ is determined through the equality

$$< Tx, y^* >_{\mathbb{C}} = < x, T^* y^* >_{\mathbb{C}} .$$

The notion of self-adjointness is now extended to this setting. A map $Z : \mathscr{X} \longrightarrow \mathscr{X}^*$ will be called **self-adjoint** or **Hermitian** if and only if $Z^* = Z$, i.e., if, for all $f, g \in \mathscr{X}$,

$$< Zf, g >_{\mathbb{C}} = < f, Zg >_{\mathbb{C}} .$$

If $\mathscr{B}$ is a basis in $\mathscr{X}$ and $\mathscr{B}^*$ is its dual basis, then the bilinear form $< Zf, g >$ can be evaluated as $([Z]_{\mathscr{B}}^{\mathscr{B}^*}[f]^{\mathscr{B}}, [g]^{\mathscr{B}})$. Here $[Z]_{\mathscr{B}}^{\mathscr{B}^*}$ is the representing matrix of Z and

$$(\xi, \eta) = \xi^* \eta := \overline{\xi}^\top \eta$$

denotes the **standard Hermitian inner product** in $\mathbb{C}^n$. One denotes by $A^* = \overline{A}^\top$ the **Hermitian adjoint** of a complex matrix $A \in \mathbb{C}^{n \times n}$. A^* is a unique matrix such that $(A\xi, \eta) = (\xi, A^* \eta)$ for all $\xi, \eta \in \mathbb{C}^n$. A matrix A is Hermitian if and only if $A^* = A$. Thus the matrix representing the Hermitian adjoint Z^* is the Hermitian adjoint A^* of $A = [Z]_{\mathscr{B}}^{\mathscr{B}^*} \in \mathbb{C}^{n \times n}$. Thus Z is Hermitian if and only if its matrix representation A is Hermitian. The map Z is called **positive definite**, denoted by $Z \succ 0$, if $< Zf, f >_{\mathbb{C}} > 0$ for all nonzero f in $\mathscr{X}$. Similarly, we write $Z \succeq 0$ whenever Z is positive semidefinite. It is easily seen that a Hermitian map Z is positive if and only if $A = [Z]_{\mathscr{B}}^{\mathscr{B}^*}$ is a positive definite Hermitian matrix.

Our starting point for the stability analysis of linear systems, induced by a complex $n \times n$ matrix A, is to derive a characterization linking the stability question with an associated eigenvalue problem.

Proposition 5.41. *1. The continuous-time system $\dot{x} = Ax$ is asymptotically stable if and only if all eigenvalues of A are in the open left half-plane $\mathbb{C}_- = \{z \in \mathbb{C} \mid \operatorname{Re} z < 0\}$.*

2. The discrete-time system $x_{t+1} = Ax_t$ is asymptotically stable if and only if all eigenvalues of A are in the open unit disc $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$.

Proof. 1. Assume that $\dot{x} = Ax$ is asymptotically stable, i.e., $\lim_{t\to\infty} e^{tA}x_0 = 0$ is satisfied for all $x_0 \in \mathbb{C}^n$. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A with associated eigenvector $v \in \mathbb{C}^n$. Then $e^{tA}v = e^{t\lambda}v$ is true for all t. Thus

$$\lim_{t\to\infty} e^{tA}v = \lim_{t\to\infty} e^{t\lambda}v = 0,$$

and therefore $\lim_{t\to\infty} e^{t\,\mathrm{Re}\,\lambda} = \lim_{t\to\infty} |e^{t\lambda}| = 0$. Thus $\lambda < 0$, and therefore the condition is necessary.

Conversely, assume that all eigenvalues of A have negative real part. Let $\lambda_1, \ldots, \lambda_r$ denote the distinct eigenvalues of A with algebraic multiplicities $n_1, \ldots, n_r$, respectively. Using the Jordan canonical form for A, one obtains the representation

$$e^{tA} = \sum_{j=1}^{r} e^{t\lambda_j} P_j(t), \tag{5.57}$$

where $P_j(t) \in \mathbb{C}^{n\times n}$ are suitable matrix polynomials in t whose entries have degrees $\leq n_j$. Thus, each entry of $e^{t\lambda_j}P_j(t)$ is of the form $e^{t\lambda_j}\pi(t)$ for a certain polynomial $\pi(t)$. But

$$\lim_{t\to\infty} |e^{t\lambda_j}\pi(t)| = \lim_{t\to\infty} e^{t\,\mathrm{Re}\,\lambda_j}|\pi(t)| = 0$$

since $\mathrm{Re}\,\lambda_j < 0$ and the exponential function grows faster than polynomials. This implies $\lim_{t\to\infty} e^{tA} = 0$, i.e., the asymptotic stability of $\dot{x} = Ax$.

2. The proof of the discrete-time case runs completely similar to the preceding case and is therefore omitted. ∎

Because of the preceding characterization, we call $\mathbb{C}_-$ and $\mathbb{D}$ the **stability domain** of $\dot{x} = Ax$ and $x_{t+1} = Ax_t$, respectively.

Definition 5.42. A scalar complex polynomial $p(z) = \sum_{i=0}^{n} a_i z^i \in \mathbb{C}[z]$ is called a **Hurwitz polynomial** (or **Schur polynomial**) if $p(z)$ has roots only in $\mathbb{C}_-$ (or only in $\mathbb{D}$).

Of course, it is by no means obvious how one can recognize whether a particular polynomial $p(z)$ is Hurwitz or Schur. For real polynomials of degree 2 it is an easy exercise to see that $z^2 + az + b$ is a Hurwitz polynomial if and only if $a > 0$ and $b > 0$. The characterization of degree 3 real Hurwitz polynomials is due to Maxwell (1868), who showed that $z^3 + az^2 + bz + c$ is Hurwitz if and only if

$$a > 0, \quad c > 0, \quad ab > c.$$

For complex polynomials, these expressions are more complicated. A full characterization of Hurwitz polynomials, expressed in terms of the positivity of a certain quadratic form whose coefficients are quadratic polynomials in the coefficients of $p(z)$, will be given later on.

Next, the classic characterization by Lyapunov (1893) of asymptotic stability in terms of linear matrix inequalities is presented.

Theorem 5.43 (Lyapunov). *Let $\mathscr{X}$ be a complex n-dimensional vector space and $A : \mathscr{X} \longrightarrow \mathscr{X}$ a linear operator. The following statements are equivalent:*

1. *The system $\dot{x} = Ax$ on $\mathscr{X}$ is asymptotically stable.*
2. *There exists a positive definite Hermitian linear operator $Q : \mathscr{X} \longrightarrow \mathscr{X}^*$ that satisfies the* ***Lyapunov inequality***

$$A^*Q + QA \prec 0.$$

Proof. Without loss of generality, one can assume that $\mathscr{X} = \mathbb{C}^n$ and $A \in \mathbb{C}^{n\times n}$. Consequently, $0 \prec Q = Q^* \in \mathbb{C}^{n\times n}$. Suppose that Q is a positive definite solution of the Lyapunov inequality. Let λ be an eigenvalue of A with associated eigenvector $v \in \mathbb{C}^n$. Then $Av = \lambda v$ and $v^*A^* = \overline{\lambda} v^*$. Thus

$$0 > v^*A^*Qv + v^*QA^*v = (\overline{\lambda} + \lambda)v^*Qv.$$

By the positive definiteness of Q, we obtain $v^*Qv > 0$, which implies $2\,\mathrm{Re}\,\lambda = \overline{\lambda} + \lambda < 0$. Thus A is asymptotically stable.

For the converse assume that $\dot{x} = Ax$ is asymptotically stable, i.e., A has only eigenvalues with negative real part. From the decomposition (5.57), it can be seen that each entry of e^{tA} is in $L^2([0,\infty),\mathbb{C})$, with $\lim_{t\to\infty} e^{tA} = 0$, and, in particular, the integral

$$Q := \int_0^\infty e^{tA^*} e^{tA} dt$$

exists. Obviously, $Q = Q^*$ is Hermitian and satisfies

$$v^*Qv = \int_0^\infty \|e^{tA}v\|^2 dt \geq 0$$

for all $v \in \mathbb{C}^n$. Thus $v^*Qv = 0$ if and only if $e^{tA}v = 0$ for all $t \geq 0$, i.e., if and only if $v = 0$. This shows that Q is positive definite. Moreover,

$$A^*Q + QA = \int_0^\infty \frac{d}{dt}(e^{tA^*} e^{tA}) dt = -I_n \prec 0.$$

This completes the proof. ■

A useful strengthening of the preceding theorem of Lyapunov is due to Snyders and M. Zakai (1970) and Wimmer (1974a).

Theorem 5.44. *Let $\mathscr{X}$ and $\mathscr{Y}$ be two complex, finite-dimensional vector spaces. Let $A : \mathscr{X} \longrightarrow \mathscr{X}$ be a linear transformation. The following assertions are equivalent:*

1. *The system $\dot{x} = Ax$ on $\mathscr{X}$ is asymptotically stable.*
2. *For a linear transformation $C : \mathscr{X} \longrightarrow \mathscr{Y}$, with (C,A) observable, the* ***Lyapunov equation***

$$A^*Q + QA = -C^*C \tag{5.58}$$

has a unique positive definite Hermitian solution $Q : \mathscr{X} \longrightarrow \mathscr{X}^*$.

3. *There exists a linear transformation $C : \mathscr{X} \longrightarrow \mathscr{Y}$, with (C,A) observable, such that the Lyapunov equation (5.58) has a positive definite Hermitian solution $Q : \mathscr{X} \longrightarrow \mathscr{X}^*$.*

Proof. Again, and without loss of generality, assume $\mathscr{X} = \mathbb{C}^n$, $\mathscr{Y} = \mathbb{C}^p$, $A \in \mathbb{C}^{n\times n}, Q = Q^* \succ 0$. Obviously, 2 implies 3.

Assume that 3 is satisfied. Let λ be an eigenvalue of A with associated eigenvector $v \in \mathbb{C}^n$. From the Lyapunov equation we get

$$2\operatorname{Re}\lambda v^*Qv = v^*A^*Qv + v^*QA^*v = -\|Cv\|^2 \leq 0.$$

Since $v^*Qv > 0$, $\operatorname{Re}\lambda \leq 0$. Suppose $\operatorname{Re}\lambda = 0$. Then $Cv = 0$, i.e., v is an eigenvector of A that is contained in the kernel of C. But, by the Hautus criterion, this contradicts observability. Therefore, each eigenvalue of A has negative real part and assertion 1 is proved.

To prove that 1 implies 2, we proceed similarly to the proof of the Lyapunov theorem. Thus, assume that A has only eigenvalues with negative real part. For $C \in \mathbb{C}^{p\times n}$, the integral

$$Q := \int_0^\infty e^{tA^*}C^*Ce^{tA}dt$$

exists and defines a Hermitian matrix. For each complex vector v this implies

$$v^*Qv = \int_0^\infty \|Ce^{tA}v\|^2 dt \geq 0,$$

and therefore $v^*Qv = 0$ if and only if $Ce^{tA}v = 0$ for all $t \geq 0$. Equivalently, $CA^kv = 0$ for all $k \in \mathbb{N}_0$. By the observability of (C,A), $v = 0$. This shows that Q is positive definite. Moreover,

$$A^*Q + QA = \int_0^\infty \frac{d}{dt}e^{tA^*}C^*Ce^{tA}dt = -C^*C.$$

This completes the proof. ∎

The preceding results by Lyapunov, Snyders-Zakai, and Wimmer have been generalized by Ostrowski and Schneider (1962) into an inertia theorem that relates the number of stable and unstable eigenvalues with the signature of Q. Their result will be proved in a slightly more general form. For a matrix $A \in \mathbb{C}^{n\times n}$, let $n_0(A)$ denote the number of eigenvalues (counting multiplicities) of A with real part zero. Similarly, let $n_\pm(A)$ denote the number of eigenvalues (counting multiplicities) of A with positive and negative real parts, respectively. The triple $\mathrm{ind}(A) = (n_0(A), n_+(A), n_-(A))$ is called the **inertia index** of A.

Theorem 5.45 (Inertia Theorem). *Let $A : \mathscr{X} \longrightarrow \mathscr{X}$ and $C : \mathscr{X} \longrightarrow \mathscr{Y}$ be linear transformations between finite-dimensional complex vector spaces. Assume that (C,A) is an observable pair. Then every Hermitian solution $Q : \mathscr{X} \longrightarrow \mathscr{X}^*$ of the* ***Lyapunov inequality***

$$A^*Q + QA + C^*C \preceq 0$$

satisfies

$$n_0(Q) = n_0(A) = 0, \quad n_+(Q) = n_-(A), \quad n_-(Q) = n_+(A). \tag{5.59}$$

Proof. First, one proves $n_0(A) = n_0(Q) = 0$. As before, one assumes, without loss of generality, that $\mathscr{X} = \mathbb{C}^n$, and so forth. Suppose that $\lambda = i\omega$ is a purely imaginary eigenvalue of A with eigenvector $v \in \mathbb{C}^n$. Multiplying v^* and v on both sides of the Lyapunov inequality, we obtain

$$0 \geq v^*(A^*Q + QA + C^*C)v = \|Cv\|^2,$$

which implies $Cv = 0$. By the observability of (C,A), $v = 0$. Thus $n_0(A) = 0$. Next, consider $v \in \mathbb{C}^n$, with $Qv = 0$. Then, by the same reasoning, we obtain $Cv = 0$. After applying a suitable unitary state-space similarity transformation, one can assume that

$$A^*Q + QA + C^*C = -\mathrm{diag}\,(I_r, 0).$$

Partition v accordingly as $v = \mathrm{col}\,(x,y)$, with $x \in \mathbb{C}^r$. Thus

$$-\|x\|^2 = v^*(A^*Q + QA + C^*C)v = 0,$$

and therefore $x = 0$. Moreover, $QAv = (A^*Q + QA + C^*C)v = -\mathrm{diag}\,(I_r, 0)v = 0$ for $v = \mathrm{col}\,(0,y)$. This implies $A(\mathrm{Ker}\, Q) \subset \mathrm{Ker}\, Q$. Since $\mathrm{Ker}\, Q \subset \mathrm{Ker}\, C$, the observability of (C,A) implies $\mathrm{Ker}\, Q = \{0\}$. Thus $n_0(Q) = 0$.

For the proof of the remaining equalities (5.59), we proceed by a simple continuity argument. Consider the convex cone of Hermitian $n \times n$ matrices

$$\mathscr{C} = \{Q = Q^* \mid A^*Q + QA + C^*C \preceq 0\}.$$

Using the observability of the pair (C,A), it was just shown that each element of $\mathscr{C}$ is invertible. Thus $\mathscr{C} \subset GL_n(\mathbb{C})$ is a convex and hence connected subset of $GL_n(\mathbb{C})$. Since the inertia index, $\mathrm{ind}(P)$, depends continuously on $P \in GL_n(\mathbb{C})$, it suffices to establish the inertia equalities (5.59) for a single element $Q_0 \in \mathscr{C}$. Since A has no purely imaginary eigenvalues, we can assume without loss of generality that (C,A) are partitioned as

$$A = \begin{pmatrix} A_+ & 0 \\ 0 & A_- \end{pmatrix}, \quad C = \begin{pmatrix} C_+ & C_- \end{pmatrix}.$$

Here the eigenvalues of A_+ are assumed to have positive real part, while the eigenvalues of A_- have negative real part. Using the Hautus criterion one sees that the observability of (C,A) is equivalent to the observability of (C_+,A_+) and (C_-,A_-). Define for real numbers $r > 0$

$$Q_- = r \int_0^\infty e^{tA_-^*} e^{tA_-} dt,$$
$$Q_+ = r \int_0^\infty e^{-tA_+^*} e^{-tA_+} dt.$$

Then $Q_\pm = Q_\pm^* \succ 0$ are positive definite and Hermitian and the Hermitian matrix

$$Q := \mathrm{diag}\,(-Q_+, Q_-)$$

satisfies

$$A^*Q + QA + C^*C = C^*C - rI.$$

Moreover, $n_+(Q) = n_-(A)$ and $n_-(Q) = n_+(A)$. Choose $r > 0$ such that $rI \succeq C^*C$. Then Q satisfies the Lyapunov inequality. ■

As an example, we examine the stability of second-order systems of the form

$$\ddot{x}(t) + (M + \Delta)\dot{x}(t) + Nx(t) = 0 \tag{5.60}$$

for $x \in \mathbb{R}^n$. The following assumptions will be made:

$$M = M^\top \succeq 0, \quad N = N^\top \succ 0,$$
$$\Delta^\top = -\Delta.$$

Consider the matrices

$$A = \begin{pmatrix} 0 & I_n \\ -N & -M - \Delta \end{pmatrix}, \quad C = \begin{pmatrix} 0 & M^{\frac{1}{2}} \end{pmatrix}.$$

Using the skew symmetry of Δ, one verifies that the positive definite symmetric matrix

$$Q = \frac{1}{2}\begin{pmatrix} N & 0 \\ 0 & I \end{pmatrix}$$

satisfies the Lyapunov equation

$$A^\top Q + QA = -C^\top C.$$

Moreover, the pair (C,A) is easily seen to be observable. Now Theorem 5.44 implies the asymptotic stability of (5.60).

2. Complex Polynomials and Hermitian Operators.

Sesquilinear forms and adjoint operators are then extended to the context of polynomial and rational models. We will discuss briefly two-variable polynomial matrices and examine naturally induced linear maps and Hermitian forms in complex polynomial models. In the preceding parts of this chapter, the connection between tensor products, linear transformations, and bilinear forms was established. When the field is taken to be the field $\mathbb{C}$ of complex numbers, this must be slightly modified in order for it to be applicable to stability questions. Using sesquilinear forms, the duality theory of polynomial models developed in Chapter 3 can be extended in a rather straightforward way.

Our analysis starts with functions in one variable. Define, for $A(z) = \sum_{j=-\infty}^{n_A} A_j z^j \in \mathbb{C}((z^{-1}))^{p\times m}$, the conjugate power series $A^* \in \mathbb{C}((z^{-1}))^{p\times m}$ by

$$A^*(z) = \overline{A(\overline{z})}^\top = \sum_{j=-\infty}^{n_A} A_j^* z^j \in \mathbb{C}((z^{-1}))^{m\times p}.$$

In $\mathbb{C}((z^{-1}))^m \times \mathbb{C}((z^{-1}))^m$, a sesquilinear form $[g,f]_{\mathbb{C}}$ is defined by

$$[g,f]_{\mathbb{C}} = \sum_{j=-\infty}^{\infty} g_{-j-1}^* f_i = (g^*(z)f(z))_{-1}, \tag{5.61}$$

where $f(z) = \sum_{j=-\infty}^{\infty} f_j z^j$, $g(z) = \sum_{j=-\infty}^{\infty} g_j z^j$, and $g^*(z) = \sum_{j=-\infty}^{\infty} g_j^* z^j$. Thus $\overline{[g,f]_{\mathbb{C}}} = [f,g]_{\mathbb{C}}$. It is clear that, because both $f(z)$ and $g(z)$ are truncated Laurent series, the sum in (5.61) is well defined, containing only a finite number of nonzero terms. Let $A(z) \in \mathbb{C}((z^{-1}))^{m\times m}$. Then $[g,Af]_{\mathbb{C}} = [A^*g,f]_{\mathbb{C}}$ for all $f(z),g(z) \in \mathbb{C}((z^{-1}))^m$. This global form is used to obtain a concrete representation of X_D^*, the dual space of the polynomial model X_D.

Proposition 5.46. *Let $D(z) \in \mathbb{C}[z]^{m\times m}$ be nonsingular. Then $D^*(z) := \overline{D(\overline{z})}^\top \in \mathbb{C}^{m\times m}[z]$ is nonsingular, and the following assertions are in force.*

1. *The dual space* X_D^* *of* X_D *can be identified with* X_{D^*} *under the nondegenerate pairing*

$$< g, f >_{\mathbb{C}} = [g, D^{-1} f]_{\mathbb{C}} = (g^*(z) D(z)^{-1} f(z))_{-1},$$

for $f(z) \in X_D$ *and* $g(z) \in X_{D^*}$.

2. *The form* $< g, f >_{\mathbb{C}}$ *is sesquilinear, i.e.,*

$$< g, f >_{\mathbb{C}} = \overline{< f, g >_{\mathbb{C}}}.$$

3. *The module structures of* X_D *and* X_{D^*} *are related through*

$$S_D^* = S_{D^*}.$$

Proof. The nondegeneracy of $< \cdot, \cdot >_{\mathbb{C}}$ follows from Theorem 3.38. Computing

$$\overline{< f, g >_{\mathbb{C}}} = \overline{(f^*(z) D^*(z)^{-1} g(z))_{-1}} = (g^*(z) D(z)^{-1} f(z))_{-1}$$

proves the second assertion. The last assertion is proved by a trivial calculation, which is omitted. ■

Let $\mathbb{C}[z,w]^{n_1 \times n_2}$ denote the $n_1 \times n_2$ complex polynomial matrices in the complex variables z and w. For $M(z,w) \in \mathbb{C}[z,w]^{n_1 \times n_2}$, one defines the conjugate polynomial $M^*(z,w) \in \mathbb{C}[z,w]^{n_2 \times n_1}$ by

$$M^*(z,w) := \overline{M(\overline{w}, \overline{z})}^{\top}.$$

A polynomial matrix $M(z,w) \in \mathbb{C}[z,w]^{n \times n}$, with

$$M(z,w) = \sum_{i,j=1}^{d} M_{ij} z^{i-1} w^{j-1},$$

will be called **Hermitian** if

$$M^*(z,w) = M(z,w).$$

It is easy to see that this is equivalent to the condition $M_{ij} = M_{ji}^*$. A Hermitian polynomial matrix $M(z,w) \in \mathbb{C}[z,w]^{n \times n}$ is called **nonnegative**, and denoted by $M \geq 0$, if and only if $\sum_{i,j} \xi_i^* M_{ij} \xi_j \geq 0$ for all $\xi_i \in \mathbb{C}^n$. This is equivalent to the matrix $(M_{ij}) \in \mathbb{C}^{dn \times dn}$ being Hermitian and positive semidefinite. Thus there exists a full column rank matrix $C \in \mathbb{C}^{dn \times k}$ with $(M_{ij}) = CC^*$. This implies that the matrix polynomial M is nonnegative if and only if there exists some $C(z) \in \mathbb{C}[z]^{n \times k}$ such that

$$M(z,w) = C(z)C^*(w).$$

In fact, the relation between C and $C(z)$ is simply $C(z) = (I_n, \ldots, z^{d-1}I_n)C$.

Next, the corresponding objects in the context of polynomial models are introduced. Let $D_1(z) \in \mathbb{C}[z]^{m\times m}$ and $D_2(z) \in \mathbb{C}[z]^{p\times p}$ be two nonsingular polynomial matrices, and let X_{D_1} and X_{D_2} be the associated polynomial models. The Kronecker product model $X_{D_2(z)\otimes D_1^\top(w)}$ is defined as the set of all complex $p \times m$ polynomial matrices $M(z,w)$ for which $D_2(z)^{-1}M(z,w)D_1(w)^{-1}$ is strictly proper in both variables. The isomorphism $X_{D_2(z)\otimes D_1^\top(w)} \simeq X_{D_2(z)} \otimes_{\mathbb{C}} X_{D_1^\top(w)}$ implies that every $M(z,w) \in X_{D_2(z)\otimes_{\mathbb{C}} D_1^\top(w)}$ has a representation, in terms of generating elements, of the form

$$M(z,w) = C_2(z)C_1^*(w),$$

with each column of $D_2(z)^{-1}C_2(z)$ and $D_1^*(w)^{-1}C_1(w)$ strictly proper. Note that

$$M^*(z,w) = C_1(z)C_2^*(w).$$

For polynomials $M(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$, we define a map $\mathbf{M} : X_{D_1} \longrightarrow X_{D_2}$, induced by $M(z,w)$, by

$$\mathbf{M}f(z) := (M(z,w)D_1^{-1}(w)f(w))_{-1} = < M^*(\cdot,\bar{z}), f >_{\mathbb{C}}, \qquad f(z) \in X_{D_1}. \tag{5.62}$$

Note that this construction parallels that in Theorem 5.34. Clearly, $\mathbf{M}$ is a linear operator that maps X_{D_1} to X_{D_2}.

Definition 5.47. A Hermitian polynomial $M(z,w) \in X_{D(z)\otimes\overline{D}(w)}$ is ***D*-positive**, denoted by $M >_D 0$, if the induced Hermitian map $\mathbf{M} : X_{D^*} \longrightarrow X_D$ is positive, that is, if the **quadratic form** $< \mathbf{M}f, f >$ is positive definite, i.e., $< \mathbf{M}f, f > > 0$ for all nonzero $f(z) \in X_{D^*}$.

Theorem 5.48. *Let $M(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$. Let the map $\mathbf{M} : X_{D_1} \longrightarrow X_{D_2}$ be defined by (5.62). Then:*

1. *The Hermitian adjoint map $\mathbf{M}^* : X_{D_2^*} \longrightarrow X_{D_1^*}$ is given, for $g(z) \in X_{D_2^*}$, by*

$$(\mathbf{M}^*g)(w) = < M^*(w,\cdot), g >_{\mathbb{C}}; \tag{5.63}$$

2. *$M(z,w) \in X_{D(z)\otimes\overline{D}(w)}$ is Hermitian if and only if $M(z,w)$ has a representation of the form*

$$M(z,w) = \sum_{i=1}^{k} \lambda_i g_i(z) g_i^*(w), \tag{5.64}$$

with $g_i(z) \in X_D$ and $\lambda_i \in \mathbb{R}$. $M(z,w)$ is Hermitian if and only if the induced map $\mathbf{M} : X_{D^*} \longrightarrow X_D$ is Hermitian;

3. *Let $D(z) \in \mathbb{C}[z]^{m\times m}$ be nonsingular, and let $M(z,w) \in X_{D(z)\otimes\overline{D}(w)}$ be Hermitian. The following conditions are equivalent:*

   a. *$M(z,w)$ is D-positive.*
   b. *There exists a basis $\{g_i(z)\}$ in X_D for which*

$$M(z,w) = \sum_{i=1}^{n} g_i(z) g_i^*(w).$$

Proof. From the identity $M(z,w) = C_2(z)C_1^*(w)$ one obtains

$$\mathbf{M}f(z) = (C_2(z)C_1^*(w)D_1(w)^{-1}f(w))_{-1},$$

and thus, for all $f \in X_{D_1}, g \in X_{D_2^*}$,

$$\begin{aligned} < g, \mathbf{M}f >_{\mathbb{C}} &= \left(g^*(z)D_2(z)^{-1}(C_2(z)C_1^*(w)D_1(w)^{-1}f(w))_{-1}\right)_{-1} \\ &= \left((g^*(z)D_2(z)^{-1}C_2(z))_{-1}C_1^*(w)D_1(w)^{-1}f(w)\right)_{-1} \\ &=< \mathbf{M}^* g, f >_{\mathbb{C}} \\ &= \left((\mathbf{M}^* g)^*(w)D_1(w)^{-1}f(w)\right)_{-1}. \end{aligned}$$

Thus

$$\begin{aligned} \mathbf{M}^* g(w) &= C_1(w)\overline{(g^*(z)D_2(z)^{-1}C_2(z))}_{-1} \\ &= \left(C_1(w)C_2^*(z)D_2^*(z)^{-1}g(z)\right)_{-1} \\ &= \left(M^*(w,z)D_2^*(z)^{-1}g(z)\right)_{-1} \\ &=< M^*(w,\cdot), g >_{\mathbb{C}}. \end{aligned}$$

This proves (5.63).

A complex matrix Q is Hermitian if and only if $Q = C\Lambda C^*$ is satisfied for a real diagonal matrix Λ and a complex matrix C. This shows that a complex polynomial $M(z,w) \in \mathbb{C}[z,w]^{m\times m}$ is Hermitian if and only if it is of the form

$$M(z,w) = C(z)\Lambda C^*(w) = \sum_{i=1}^{k} \lambda_i g_i(z) g_i^*(w)$$

for suitable real numbers $\lambda_1, \dots, \lambda_k$ and complex polynomial vectors $g_i(z) \in \mathbb{C}[z]^m$. Here $g_i(z)$ denotes the ith column of $C(z)$. Also, $M(z,w) \in X_{D(z)\otimes\overline{D}(w)}$ if and only if $g_i \in X_D$ for all i. This proves (5.64). For $f \in X_{D^*}$, one has

$$\mathbf{M}f = \left(M(z,w)D^*(w)^{-1}f(w)\right)_{-1}$$
$$\mathbf{M}^*f = \left(M^*(w,z)D^*(z)^{-1}f(z)\right)_{-1}.$$

Thus $\mathbf{M} = \mathbf{M}^*$ if and only if $M^*(z,w) = M(z,w)$. This completes the proof of the second assertion.

To prove the last assertion, one computes the inner product $< \mathbf{M}f, f >_{\mathbb{C}}$ for each $f \in X_{D^*}$. Since $M(z,w)$ is Hermitian, it has the representation $M(z,w) = C(z)\Lambda C^*(w)$ for a real diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ and $D(z)^{-1}C(z)$ strictly proper. Without loss of generality, one can assume that the columns $c_1(z), \ldots, c_n(z)$ of $C(z)$ form a basis of X_D. Thus the quadratic form

$$\begin{aligned} < \mathbf{M}f, f >_{\mathbb{C}} &= \left(f^*(z)D(z)^{-1}(C(z)\Lambda C^*(w)D^*(w)^{-1}f(w))_{-1}\right)_{-1} \\ &= \left(f^*(z)D(z)^{-1}C(z)\right)_{-1}\Lambda\left(C^*(w)D^*(w)^{-1}f(w)\right)_{-1} \\ &= \sum_{i=1}^{n} \lambda_i | < f, c_i >_{\mathbb{C}} |^2 \end{aligned}$$

is positive definite on X_{D^*} if and only if $\lambda_1 > 0, \ldots, \lambda_n > 0$. Thus the elements $g_i(z) = \sqrt{\lambda_i}c_i(z)$, $i = 1, \ldots, n$, define a basis of X_D, with

$$M(z,w) = \sum_{i=1}^{n} g_i(z)g_i^*(w)$$

and

$$< \mathbf{M}f, f >_{\mathbb{C}} = \sum_{i=1}^{n} | < f, g_i >_{\mathbb{C}} |^2.$$ ■

In Theorem 5.48, the map induced by a two-variable polynomial matrix $M(z,w) \in X_{D_2(z)\otimes D_1^\top(w)}$ was examined. This restriction is unnecessary because one can use the projection $\pi_{D_2(z)\otimes D_1^\top(w)}$ in $\mathbb{C}[z,w]^{p\times m}$, as defined in (5.20). Thus, assume that $M(z,w) \in \mathbb{C}[z,w]^{p\times m}$. Let $D_1(z) \in \mathbb{C}[z]^{m\times m}$ and $D_2(z) \in \mathbb{C}[z]^{p\times p}$ be two nonsingular polynomial matrices, and let X_{D_1} and X_{D_2} be the associated polynomial models. Since $\mathrm{Im}\,\pi_{D_2(z)\otimes D_1^\top(w)} = X_{D_2(z)\otimes_{\mathbb{C}} D_1^\top(w)}$, one can define the induced map $\mathbf{M} : X_{D_1} \longrightarrow X_{D_2}$ by

$$(\mathbf{M}f)(z) = \left(\pi_{D_2(z)\otimes D_1^\top(w)} M(z,w)D_1(w)^{-1}f(w)\right)_{-1}. \tag{5.65}$$

In view of (5.24), the induced map satisfies $\mathbf{M} = 0$ if and only if there exist polynomial matrices $M_i(z,w)$ such that $M(z,w) = D_2(z)M_1(z,w) + M_2(z,w)D_1(w)$.

3. Stability of Higher-Order Equations
We proceed now to establish stability criteria for complex polynomial matrices. This is done by reduction to the Lyapunov theorem. For simplicity we focus on continuous-time systems; discrete-time systems can be handled similarly. Recall that a complex matrix A is called **stable** if all its eigenvalues lie in the open left half-plane. Equivalently, a matrix A is stable if and only if its characteristic polynomial is stable. One extends this definition to nonsingular polynomial matrices by saying that λ is an **eigenvalue of a nonsingular polynomial matrix** $D(z)$ if there exists a nonzero vector ξ in $\operatorname{Ker} D(\lambda)$. In view of Theorem 3.30, λ is an eigenvalue of $D(z)$ if and only if it is an eigenvalue of S_D. A nonsingular polynomial matrix $D(z)$ is a **stable polynomial matrix** if all its eigenvalues lie in the open left half-plane. Note that a matrix A is stable if and only if the polynomial matrix $zI - A$ is stable. With the analysis of stability our present goal, it is important to introduce symmetry with respect to the imaginary axis. For a polynomial matrix $P(z) \in \mathbb{C}[z]^{m\times m}$, one defines its **parahermitian adjoint** $P_*(z)$ by

$$P_*(z) := \overline{P(-\overline{z})}^{\top} = P^*(-z).$$

One says that $P(z)$ is **parahermitian** if $P_*(z) = P(z)$. Clearly, $P(z)$ is stable if and only if $P_*(z)$ is antistable, i.e., it has all its eigenvalues in the open right half-plane.

Our principal theorem is stated next.

Theorem 5.49. *Let $P(z) \in \mathbb{C}[z]^{m\times m}$ be a nonsingular polynomial matrix. Then $P(z)$ is stable if and only if, for polynomial matrices $R(z)$ with $P(z)$ and $R(z)$ right coprime and $R(z)P(z)^{-1}$ proper, there exists a solution $Q(z)$ of the polynomial Sylvester equation*

$$P_*(z)Q(z) + Q_*(z)P(z) = R_*(z)R(z) \tag{5.66}$$

such that the quadratic form, induced in X_P by

$$V(z,w) = \frac{P^*(z)Q(w) + Q^*(z)P(w) - R^*(z)R(w)}{z+w}, \tag{5.67}$$

is positive definite.

Proof. Assume $P(z)$ is stable and $R(z)$ is right coprime with $P(z)$. The coprimeness condition implies (Theorem 4.28) that the pair (C,A) defined, in the state space X_P, by the shift realization (4.23) is observable. Explicitly, the shift realization of the proper transfer function $R(z)P(z)^{-1}$, with $R(z)P(z)^{-1} - D$ being strictly proper, is given as $A : X_P \longrightarrow X_P$, $C : X_P \longrightarrow \mathbb{C}^p$,

$$\begin{aligned} A &= S_P, \\ Cf &= \left(R(z)P(z)^{-1}f(z)\right)_{-1}, \quad f \in X_P. \end{aligned} \tag{5.68}$$

Note that $Cf = \left((R(z)P(z)^{-1} - D)f(z)\right)_{-1}$, which implies that C does not depend on the constant term D. A straightforward computation shows that the Hermitian adjoint of C is the linear operator $C^* : \mathbb{C}^p \longrightarrow X_{P^*}$, defined as $C^*v = R^*(z) - P^*(z)D^*v$ for $v \in \mathbb{C}^p$. In fact, for all $v \in \mathbb{C}^p, f \in X_P$,

$$< C^*v, f >_{\mathbb{C}} = \left((C^*v)^*(z)P(z)^{-1}f(z)\right)_{-1} =< v, Cf >_{\mathbb{C}}$$
$$= v^*\left((R(z) - DP(z))P(z)^{-1}f(z)\right)_{-1},$$

which is equivalent to $C^* = R^*(z) - P^*(z)D^*$.

The stability of $P(z)$ implies that there exists a solution to equation (5.66). This is shown next. Because $P(z)$ is stable, so is $p(z) = \det P(z)$, and the scalar polynomials $p(z)$ and $p_*(z) := \overline{p}(-z)$ are coprime. Therefore, the polynomial Sylvester equation

$$K(z)p(z) + p_*(z)L(z) = R^*(-z)R(z)$$

is solvable. Moreover, the solution is unique if one assumes $L(z)$ is reduced modulo $p(z)$ and $K(z)$ modulo $p_*(z)$. By a symmetry argument, $K(z) = L_*(z)$, and hence

$$L_*(z)p(z) + p_*(z)L(z) = R_*(z)R(z).$$

Using Cramer's rule, i.e., $p(z)I = \operatorname{adj} P(z)P(z)$, and defining $Q(z) := \operatorname{adj} P_*(z)L(z)$, with $Q_*(z) = L_*(z)\operatorname{adj} P(z)$, the equality $Q_*(z)P(z) + P_*(z)Q(z) = R_*(z)R(z)$ follows.

Thus $Q(z)$ solves the polynomial Sylvester equation (5.66). As a consequence of (5.66), $V(z,w)$ defined by (5.67) is a polynomial matrix in two variables. Let

$$M(z,w) = (z+w)V(z,w) = P^*(z)Q(w) + Q^*(z)P(w) - R^*(z)R(w).$$

Clearly, both $V(z,w)$ and $M(z,w)$ are Hermitian. Moreover,

$$\pi_{P^*(z)\otimes P^\top(w)}M(z,w) = -(R^*(z) - P^*(z)D^*)R(w).$$

The polynomial matrix $M(z,w)$ induces a Hermitian linear operator

$$\mathbf{M} : X_P \longrightarrow X_{P^*}, \quad \mathbf{M}f = \left(\pi_{P^*(z)\otimes P^\top(w)}M(z,w)P(w)^{-1}f(w)\right)_{-1}$$

by equation (5.65). Using (5.68) we obtain

$$\mathbf{M}f = -(R^*(z) - P^*(z)D^*)\left(R(w)P(w)^{-1}f(w)\right)_{-1} = -(R^*(z) - P^*(z)D^*)Cf = -C^*Cf.$$

Thus

$$\mathbf{M} = -C^*C. \tag{5.69}$$

Similarly, using Proposition 5.32, the polynomial $V(z,w)$ induces a Hermitian linear operator $\mathbf{V} : X_P \longrightarrow X_{P^*}$ defined by

$$\mathbf{V}f = \left(\pi_{P^*(z)\otimes P^\top(w)} V(z,w)P(w)^{-1}f(w)\right)_{-1}.$$

Since the adjoint S_P^* coincides with the shift S_{P^*} on X_{P^*}, one obtains, for all $f \in X_P$,

$$\begin{aligned}(S_{P^*}\mathbf{V} + \mathbf{V}S_P)f &= \left(\pi_{P^*(z)\otimes P^\top(w)}(zV(z,w) + V(z,w)w)P(w)^{-1}f(w)\right)_{-1} \\ &= \left(\pi_{P^*(z)\otimes P^\top(w)} M(z,w)P(w)^{-1}f(w)\right)_{-1} = \mathbf{M}f.\end{aligned}$$

This proves that

$$S_{P^*}\mathbf{V} + \mathbf{V}S_P = \mathbf{M}.$$

Since $A = S_P$, and using (5.69), one concludes that $\mathbf{V}$ satisfies the Lyapunov equation

$$A^*\mathbf{V} + \mathbf{V}A = -C^*C.$$

Since P is stable, the shift operator $A = S_P$ has all its eigenvalues in the open left half-plane. By Theorem 5.44, the quadratic form $< \mathbf{V}f, f >$ is positive definite, or, equivalently, $\mathbf{V} >_P 0$.

Conversely, assume $Q(z)$ is a solution of the polynomial Sylvester equation (5.66) and the quadratic form $< \mathbf{V}f, f >$ induced in X_P by $V(z,w)$, as defined in (5.67), is positive definite. Since $S_{P_*} = S_P^*$, as before, the following Lyapunov equation is satisfied:

$$S_P^*\mathbf{V} + \mathbf{V}S_P = -C^*C.$$

Applying Theorem 5.44, one concludes that S_P is stable and, hence, by Theorem 3.30, that $P(z)$ is, too. ∎

Using the same technique, one can derive a higher-order analog of the Ostrowski and Schneider (1962) inertia theorem.

Theorem 5.50 (Polynomial Inertia Theorem). *Let $G(z) = R(z)P(z)^{-1}$ be a proper complex rational matrix function that is right coprime. Suppose that the polynomial matrix $P(z)$ has n_0 eigenvalues with real part zero, n_+ eigenvalues with positive real part, and n_- eigenvalues with negative real part, all counted with* *multiplicities. Let $Q(z)$ be a solution of the polynomial Sylvester equation*

$$P_*(z)Q(z)+Q_*(z)=R_*(z)R(z),$$

and let

$$V(z,w)=\frac{P^*(z)Q(w)+Q^*(z)P(w)-R^*(z)R(w)}{z+w}$$

denote the associated Hermitian polynomial. Then the Hermitian operator $\mathbf{V}$ *in* X_P *is invertible with exactly* n_+ *negative and* n_- *positive eigenvalues. Moreover,* $n_0=0$.

Proof. Following the notation in the proof of Theorem 5.49, one obtains

$$A^*\mathbf{V}+\mathbf{V}A=-C^*C$$

for the shift realization (A,C) of $R(z)P(z)^{-1}$. By the right coprimeness of $P(z),R(z)$, the pair (C,A) is observable. Thus the result follows from the inertia theorem 5.45. ∎

4. Classical Stability Criteria

The results of the previous section can be used for an easy derivation of some of the classical stability criteria for real and complex scalar polynomials. To do this, it will be useful to compute the matrix representation of the Hermitian operator $\mathbf{V}$ induced by a scalar two-variable polynomial $V(z,w)$. Our derivation of the classical stability criteria is nonstandard insofar as they are deduced from the Lyapunov stability criteria rather than from using winding number arguments and the Cauchy index.

Let

$$V(z,w)=\sum_{i,j=1}^{n} a_{ij}z^{i-1}w^{j-1}$$

denote a Hermitian scalar polynomial, and let $p(z)=\sum_{j=0}^{n}p_jz^j, p_n=1$, denote a monic complex polynomial. Note that

$$X_{\overline{p}(z)\otimes p(w)}=\{\sum_{i,j=1}^{n} a_{ij}z^{i-1}w^{j-1} \mid a_{ij}\in\mathbb{C}\}$$

and therefore each polynomial in z,w of degree $<n$ is contained in $X_{\overline{p}(z)\otimes p(w)}$. In particular, $V(z,w)\in X_{\overline{p}(z)\otimes p(w)}$. Recall next the construction of the standard basis and control basis on scalar polynomial models. Thus, let

$$\mathscr{B}_{\mathrm{st}}=\{1,z,\ldots,z^{n-1}\}$$

denote the **standard basis** of $X_{\overline{p}}$ and

$$\mathscr{B}_{\text{co}} = \{e_1(z), \dots, e_n(z)\}$$

with

$$e_i(z) = z^{n-i} + p_{n-1}z^{n-i-1} + \cdots + p_i = \pi_+(z^{-i}p(z))$$

denote the **control basis** of X_p. The computation

$$\begin{aligned}\left(\frac{z^{k-1}e_\ell(z)}{p(z)}\right)_{-1} &= \left(\frac{z^{k-1}\pi_+(z^{-\ell}p(z))}{p(z)}\right)_{-1} \\ &= \left(\frac{z^{k-1}(z^{-\ell}p(z))}{p(z)}\right)_{-1} = \left(z^{k-\ell-1}\right)_{-1} = \delta_{k\ell}\end{aligned}$$

shows that $\mathscr{B}_{\text{co}}$ is the dual basis to $\mathscr{B}_{\text{st}} = \{1, z, \dots, z^{n-1}\}$.

To obtain a matrix representation of the operator $\mathbf{V} : X_p \longrightarrow X_{\overline{p}}$ with respect to the control basis and standard basis on X_p and X_{p^*}, respectively, one computes, using the fact that the product of two strictly proper functions has zero residue,

$$\begin{aligned}\left(V(z,w)p(w)^{-1}e_j(w)\right)_{-1} &= \sum_{r,s=1}^{n} a_{rs}z^{r-1}\left(w^{s-1}p(w)^{-1}\pi_+(w^{-j}p(w))\right)_{-1} \\ &= \sum_{r,s=1}^{n} a_{rs}z^{r-1}\left(w^{s-1}p(w)^{-1}(w^{-j}p(w))\right)_{-1} \\ &= \sum_{r,s=1}^{n} a_{rs}z^{r-1}\left(w^{s-j-1}\right)_{-1} \\ &= \sum_{r=1}^{n} a_{rj}z^{r-1}.\end{aligned}$$

Thus the matrix representation of $\mathbf{V}$ is given by the coefficients of $V(z,w)$, that is, $[\mathbf{V}]^{\text{st}}_{\text{co}} = (a_{rs})$.

Definition 5.51. Let $p(z) = z^n + p_{n-1}z^{n-1} + \cdots + p_0 \in \mathbb{C}[z]$ be a complex polynomial with Hermitian adjoint $\overline{p}(z) = z^n + \overline{p}_{n-1}z^{n-1} + \cdots + \overline{p}_0 \in \mathbb{C}[z]$. The **Hermite–Fujiwara form** is defined as a Hermitian form with generating function

$$\frac{\overline{p}(z)p(w) - p(-z)\overline{p}(-w)}{z+w} = \sum_{i,j=1}^{n} h_{ij}z^{i-1}w^{j-1},$$

and the **Hermite–Fujiwara matrix** is the Hermitian matrix

$$\mathbf{H}_n(p) = (h_{ij}) \in \mathbb{C}^{n\times n}.$$

As an example, the Hermite–Fujiwara matrix for a complex polynomial $z^2 + p_1 z + p_0$ of degree two is computed. Thus

$$\mathbf{H}_2(p) = 2\begin{pmatrix} \mathrm{Re}\,(\overline{p}_0 p_1) & -i\mathrm{Im}\, p_0 \\ i\mathrm{Im}\, p_0 & \mathrm{Re}\, p_1 \end{pmatrix},$$

which is positive definite if and only if

$$\mathrm{Re}(p_1) > 0, \quad \mathrm{Re}(p_1)\mathrm{Re}(\overline{p}_0 p_1) - \mathrm{Im}^2(p_0) > 0.$$

From the polynomial inertia theorem the following root location result can be deduced.

Theorem 5.52 (Hermite). *Let $p(z) \in \mathbb{C}[z]$ be a monic complex polynomial of degree n that is coprime with $\overline{p}(-z)$. Let n_+ and n_- denote the number of roots of $p(z)$ that are located in the open right half-plane and open left half-plane, respectively. Let $\mathbf{H}_n(p)$ denote the Hermite–Fujiwara matrix of $p(z)$. Then $\mathbf{H}_n(p)$ is invertible with exactly n_- positive eigenvalues and n_+ negative eigenvalues.*

Proof. Defining $r(z) = \overline{p}(-z)$ and $q(z) = \frac{1}{2}p(z)$, the polynomials $r(z)$ and $p(z)$ are coprime, and $q(z)$ solves the polynomial Sylvester equation

$$p(z)q_*(z) + q(z)p_*(z) = r_*(z)r(z).$$

Thus

$$V(z,w) := \frac{p^*(z)q(w) + q^*(z)p(w) - r^*(z)r(w)}{z+w} = \frac{\overline{p}(z)p(w) - p(-z)\overline{p}(-w)}{z+w}$$

coincides with the Hermite–Fujiwara form. Since $\{1, z, \dots, z^{n-1}\}$ is a basis of the polynomial model X_p, the matrix representation of $\mathbf{V}$ on X_p is given by $\mathbf{H}_n(p)$. The result follows from Theorem 5.50. ■

Theorem 5.52 can be generalized by omitting the coprimeness assumption of $p(z)$ and $\overline{p}(-z)$. The result is strong enough for the characterization of asymptotic stability.

Theorem 5.53. *A necessary and sufficient condition for a complex polynomial $p(z)$ to be a Hurwitz polynomial is that the Hermite–Fujiwara form*

$$\frac{\overline{p}(z)p(w) - p(-z)\overline{p}(-w)}{z+w}$$

must be positive definite on X_p or, equivalently, that the Hermite–Fujiwara matrix $\mathbf{H}_n(p)$ *must be positive definite.*

Proof. Assuming that $p(z)$ is a Hurwitz polynomial implies that $p(z)$ and $\overline{p}(-z)$ do not have common roots and, hence, are coprime. By Theorem 5.52, the Hermite–Fujiwara matrix $\mathbf{H}_n(p)$ is positive definite. Conversely, assume that $\mathbf{H}_n(p)$ is positive definite or, equivalently, that the Hermite–Fujiwara form

$$\frac{\overline{p}(z)p(w) - p(-z)\overline{p}(-w)}{z+w} \tag{5.70}$$

is positive definite on X_p. The change of variable $w = -\zeta$ transforms (5.70) into the **Bezoutian form**

$$\frac{\overline{p}(z)p(-\zeta) - p(-z)\overline{p}(\zeta)}{z-\zeta}$$

of $\overline{p}(z)$ and $p_*(z) = p(-z)$. This shows that the **Bezoutian matrix** $B(\overline{p}, p_*)$ of $\overline{p}$ and p_* is equal to the product

$$B(\overline{p}, p_*) = \mathbf{H}_n(p)S$$

of the Hermite–Fujiwara matrix with the invertible matrix

$$S = \operatorname{diag}(1, -1, \cdots, (-1)^{n-1}).$$

Thus $B(\overline{p}, p_*)$ is invertible and Theorem 5.40 implies that $\overline{p}$ and p_* are coprime. Equivalently, p and $\overline{p}_*$ are coprime. Theorem 5.52 is now applied to infer the stability of $p(z)$. ∎

In the case of real polynomials, the Hermite–Fujiwara form admits a further reduction. To this end, the even and odd parts $p_+(z)$ and $p_-(z)$ of a real polynomial $p(z) = \sum_{j\geq 0} p_j z^j$ are introduced. These are defined as the polynomials

$$p_+(z) = \sum_{j\geq 0} p_{2j} z^j, \quad p_-(z) = \sum_{j\geq 0} p_{2j+1} z^j.$$

Thus

$$p(z) = p_+(z^2) + zp_-(z^2), \quad p_*(z) = p_+(z^2) - zp_-(z^2). \tag{5.71}$$

In the next proposition, it will be shown that both the Bezoutian $B(p, p_*)$ and the Hermite–Fujiwara forms have direct sum decompositions that are useful for reducing the computational complexity of stability analysis.

Proposition 5.54. *Let $p(z)$ be a real polynomial.*

1. *The following isomorphisms of quadratic forms are valid: For the Hermite–Fujiwara form $H(p)$ one has*
$$H(p) \simeq 2B(zp_-, p_+) \oplus 2B(p_+, p_-),$$
whereas for the Bezoutian $B(p, p_*)$,
$$B(p, p_*) \simeq 2B(zp_-, p_+) \oplus 2B(p_-, p_+).$$
2. *The Hermite–Fujiwara form is positive definite if and only if the two Bezoutians $B(q_+, q_-)$ and $B(zq_-, q_+)$ are positive definite.*

Proof. 1. The polynomial $p(z)$ being real implies $p(z) = \overline{p}(z)$. From (5.71) it follows that $p(-z) = p_+(z^2) - zp_-(z^2)$. We compute

$$\frac{p(z)p(w) - p(-z)p(-w)}{(z+w)} = 2\frac{zp_-(z^2)p_+(w^2) + p_+(z^2)wp_-(w^2)}{z+w}$$
$$= 2\frac{z^2p_-(z^2)p_+(w^2) - p_+(z^2)w^2p_-(w^2)}{z^2-w^2} - 2zw\frac{p_-(z^2)p_+(w^2) - p_+(z^2)p_-(w^2)}{z^2-w^2}.$$

The first summand contains only even terms, while the second contains only odd ones. This proves the first statement. By a change of variable $w = -\zeta$, the Hermite–Fujiwara form transforms into the Bezoutian of $\overline{p}(z)$ and $p(-z)$. However, this change of variable affects only the terms in

$$2zw\frac{p_-(z^2)p_+(w^2) - p_+(z^2)p_-(w^2)}{z^2-w^2},$$

and this by a change of sign.

2. Follows from the direct sum representation of the Hermite–Fujiwara form.

■

The following classical result is obtained as a direct corollary of this.

Theorem 5.55. *Let $p(z)$ be a monic real polynomial of degree n. The following statements are equivalent:*

(i) $p(z)$ is a Hurwitz polynomial.
(ii) The Hermite–Fujiwara matrix $\mathbf{H}_n(p)$ is positive definite.
(iii) The two Bezoutian matrices $B(p_+, p_-)$ and $B(zp_-, p_+)$ are positive definite.

5.6 Exercises

1. Compute the tensor product $\mathbb{Z}_2 \otimes_{\mathbb{Z}} \mathbb{Z}_3$. For which pairs of integers $m, n \in \mathbb{N}$ is $\mathbb{Z}_m \otimes_{\mathbb{Z}} \mathbb{Z}_n = \{0\}$? What is $\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q}$?
2. Prove that $m \otimes n = 0$ is valid in the tensor product $M \otimes_R N$ if and only if every bilinear map $B : M \times N \longrightarrow P$ vanishes at (m, n).

3. For ideals I and J in a commutative ring R there exists a unique R-linear map
$$R/I \otimes_R R/J \simeq R/(I+J)$$
satisfying $(x+I)\otimes(y+J) \mapsto xy+(I+J)$.
4. Let $M = R^n$ be a free module with $n \geq 2$ and $\{e_1,\ldots,e_n\}$ the standard basis. Check that $e_1 \otimes e_1 + e_2 \otimes e_2$ is not an elementary tensor in $M \otimes_R M$, i.e., that there exists no $v \in M$ with $e_1 \otimes e_1 + e_2 \otimes e_2 = v \otimes v$.
5. Prove that the tensor product $M \otimes_R N$ of torsion modules M and N over a commutative ring is a torsion module. Prove that the algebraic dual $M' = \{0\}$ for all torsion modules M. What happens if R has zero divisors?
6. Prove the R-module isomorphism $M \otimes_R N \simeq N \otimes_R M$.
7. Prove that the tensor product $f \otimes g$ of two surjective R-module homomorphisms $f : M_1 \longrightarrow N_1$ and $g : M_2 \longrightarrow N_2$ is surjective. Is this also true if surjectivity is replaced by injectivity?
8. Explain why the submodule $\{\begin{pmatrix} a & b \\ b & c \end{pmatrix} | a,b,c \in \mathbb{F}[z]\} \subset \mathbb{F}[z]^{2\times 2}$ does not have a representation of the form $D_1(z)\mathbb{F}[z]^{2\times 2} + \mathbb{F}[z]^{2\times 2}D_2(z)$.
9. Let $D_1(z) \in \mathbb{F}[z]^{m\times m}$ and $D_2(z) \in \mathbb{F}[z]^{p\times p}$ be nonsingular polynomial matrices with invariant factors $d_1,\ldots,d_m$ and $e_1,\ldots,e_p$, respectively. Let $d_i \wedge e_j$ denote the greatest common divisor of the polynomials $d_i(z)$ and $e_j(z)$. Prove the dimension formula
$$\dim_{\mathbb{F}} \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}) = \sum_{i,j} \deg(d_i \wedge e_j).$$
Deduce
$$\dim_{\mathbb{F}} \mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_1}) = \sum_{i=1}^{m} (2i-1)\deg d_i.$$
10. Let $D_1(z) \in \mathbb{F}[z]^{m\times m}$ and $D_2(z) \in \mathbb{F}[z]^{p\times p}$ be nonsingular. Show that $\mathrm{Hom}_{\mathbb{F}[z]}(X_{D_1}, X_{D_2}) = \{0\}$ if and only if $\det D_1(z)$ and $\det D_2(z)$ are coprime.
11. Let $D_1(z) \in \mathbb{F}[z]^{m\times m}$ and $D_2(z) \in \mathbb{F}[z]^{m\times m}$ be nonsingular polynomial matrices with determinants $d_i(z) = \det D_i(z)$. Show that
$$X_{D_1(z)\otimes D_2(z)} = (D_1(z) \otimes I)X_{I\otimes D_2(z)} \oplus X_{D_1(z)\otimes I}(I \otimes D_2(z))$$
is true if and only if $d_1(z)$ and $d_2(z)$ are coprime.
12. Let A be a linear transformation in $\mathbb{F}^n$ with invariant factors $d_1,\ldots,d_n$. Let $\mathscr{C}(A)$ denote the centralizer of A, i.e., the set of all $Z \in \mathbb{F}^{n\times n}$, with $ZA = AZ$. Prove the following:
   a. Show $\dim_{\mathbb{F}} \mathscr{C}(A) = n$ if and only if A is cyclic.
   b. Show $\dim_{\mathbb{F}} \mathscr{C}(A) = n^2$ if and only if A is scalar, i.e., $A = \alpha I$ for some $\alpha \in \mathbb{F}$.

c. Let A and B be linear transformations. Show that there exist no nontrivial maps intertwining them if and only if the minimal polynomials or, equivalently, the characteristic polynomials of A and B are coprime.

13. Show that the discrete-time system $x_{t+1} = Ax_t$ on $\mathbb{C}^n$ is asymptotically stable if and only if for each positive definite Hermitian matrix $Q = Q^*$ there exists a unique positive definite Hermitian solution $P = P^*$ of the Stein equation

$$A^*PA - P = -Q.$$

14. Prove that a real monic polynomial $p(z)$ of degree n is a Schur polynomial if and only if

$$(z-1)^n p(\frac{z+1}{z-1})$$

is a Hurwitz polynomial.

15. Prove that every real Hurwitz polynomial $p(z) = z^n + p_{n-1}z^{n-1} + \cdots + p_0$ satisfies $p_0 > 0, p_1 > 0, \ldots, p_{n-1} > 0$.
16. Prove that $\dot{x} = Ax$ is asymptotically stable for the tridiagonal matrix

$$A = \begin{pmatrix} -\frac{1}{2}a_1^2 & -a_2 & \ldots & 0 \\ a_2 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_n \\ 0 & \ldots & a_n & 0 \end{pmatrix}$$

if $a_i \neq 0$ for $i = 1, \ldots, n$.

17. Assume that M and N are real symmetric $n \times n$ matrices that are positive definite. Prove that the second-order system

$$\ddot{x}(t) + M\dot{x}(t) + Nx(t) = 0$$

is asymptotically stable, i.e., $\lim_{t\to\infty} x(t) = 0$ and $\lim_{t\to\infty} \dot{x}(t) = 0$ are true for all solutions.

5.7 Notes and References

Our exposition of the basic theory of tensor products of modules and quotient modules follows Lang (1965) and Hungerford (1974). In the paper by Helmke and Fuhrmann (1998), tensored polynomial and rational models were introduced to describe tangent spaces of manifolds of rational functions. The systematic study of tensor products of functional models is due to Fuhrmann and Helmke (2010) and is continued in Fuhrmann (2010a). The tensor products of polynomial models lead to a polynomial approach to the Sylvester and Stein equations and clarifies the role of Bezoutians in representing solutions. The polynomial approach to Lyapunov equations is due to Willems and Fuhrmann (1992). For matrix versions of these equations see also de Souza and Bhattacharyya (1981) and Heinig and Rost (1984). The study of Bezoutians is old and dates back to the nineteenth century, with important contributions by Cayley, Jacobi, and Sylvester.

The problem of the stability of a linear (control) system was one of the first problems of the area of control theory. The interest in stability analysis is usually traced to J.C. Maxwell's theory of governors Maxwell (1868). However, the problem of root location of polynomials has a longer history. Since, with the work of Galois and Abel, exact determination of zeros of polynomials was proved to be impossible, interest shifted to the problem of localizing the zeros in some region of the complex plane. The unit disc and the major half-planes were the regions of greatest interest. The problem of root location was already solved by Hermite (1856). But in this connection the work of Routh (1877) turned out to be important because of the efficiency of the computational algorithm. In the same way, the work of Hurwitz (1895) was significant for its connection to topological problems. For a derivation of algebraic stability criteria that is close to the spirit of the chapter we refer the reader to Fuhrmann (1982).

In a major contribution to the subject, Lyapunov (1893) offered a completely different approach based on energy considerations. In the linear case, the Lyapunov theory reduces the study of the stability of a system of first-order homogeneous constant coefficient differential equations to the positive definiteness of the solution of the celebrated Lyapunov equation. This reduction is generally attributed to Gantmacher (1959). Our approach to the stability problem of higher-order systems of differential equations is based on a strengthened form of the Lyapunov equation, given in Theorem 5.44, replacing positive definiteness by a reachability, or observability, assumption. Considering Lyapunov equations in this form is due to Snyders and M. Zakai (1970) and Wimmer (1974a). Our approach to this reduction is achieved via the use of polynomial model theory and tensor algebra. The polynomial matrix analog of the Lyapunov equation is identified, and, with a solution to this equation, a two-variable polynomial matrix is constructed. In turn, this polynomial matrix induces an operator between two polynomial models. In the special case of symmetry, this map induces a quadratic form on a polynomial model, which leads to the required reduction.

The classic paper by Krein and Naimark (1936), is an excellent source for much of the older work on root location; however, strangely, no mention of Bezoutians is made there. The study of scalar Bezoutians goes back to Cayley. Multivariable Bezoutians were introduced by Anderson and Jury (1976) and used to derive rank formulas for the McMillan degree and stability test for multivariable linear systems. Their connection to tensor products and homomorphisms of polynomial models is central to our approach. Theorem 5.49 and its application to the stability analysis of higher-order equations are due to Willems and Fuhrmann (1992). The dimension formula (5.37) appears in Gantmacher (1959) and is attributed to Shoda (1929).