Chapter 5
The Shift Operator

5.1 Introduction

We now turn our attention to the study of a special class of transformations, namely shift operators. These will turn out later to serve as models for all linear transformations, in the sense that every linear transformation is similar to a shift operator.

5.2 Basic Properties

We introduce now an extremely important class of linear transformation that will play a central role in the analysis of the structure of linear transformations. Recall that for a nonzero polynomial $q(z)$, we denote by $\pi_q f$ the remainder of the polynomial $f(z)$ after division by $q(z)$. Clearly, π_q is a projection operator in $\mathbb{F}[z]$. We note that given a nonzero polynomial $q(z)$, any $f(z) \in \mathbb{F}[z]$ has a unique representation of the form

$$f(z) = a(z)q(z) + r(z), \tag{5.1}$$

with $\deg r < \deg q$. The remainder, $r = \pi_q f$, can be written in another form based on the direct sum representation (2.7), namely $\mathbb{F}((z^{-1})) = \mathbb{F}[z] \oplus z^{-1}\mathbb{F}[[z^{-1}]]$. Let π_+, π_- be defined by (1.23), i.e., the projections of $\mathbb{F}((z^{-1}))$ on $\mathbb{F}[z]$ and $z^{-1}\mathbb{F}[[z^{-1}]]$ respectively, which correspond to the above direct sum decomposition. From the representation (5.1) of $f(z)$, we have $q(z)^{-1}f(z) = a(z) + q(z)^{-1}r(z)$. Applying the projection π_-, we have $\pi_- q^{-1} f = \pi_- q^{-1} r = q^{-1} r$, which implies an important alternative representation for the remainder, namely

$$\pi_q f = q\pi_- q^{-1} f. \tag{5.2}$$

P.A. Fuhrmann, *A Polynomial Approach to Linear Algebra*, Universitext,
DOI 10.1007/978-1-4614-0338-8_5, © Springer Science+Business Media, LLC 2012

The importance of equation (5.2) stems from the fact that it easily extends to the case of polynomial vectors.

Proposition 5.1. *Let $q(z)$ be a monic polynomial in $\mathbb{F}[z]$ and let $\pi_q : \mathbb{F}[z] \longrightarrow \mathbb{F}[z]$ be the projection map defined in (5.2). Then we have*

$$\operatorname{Ker} \pi_q = q\mathbb{F}[z] \tag{5.3}$$

and the direct sum

$$\mathbb{F}[z] = X_q \oplus q\mathbb{F}[z]. \tag{5.4}$$

Defining the set X_q by

$$X_q = \operatorname{Im} \pi_q = \{\pi_q f | f(z) \in \mathbb{F}[z]\}, \tag{5.5}$$

we have the isomorphism

$$X_q \simeq \mathbb{F}[z]/q\mathbb{F}[z]. \tag{5.6}$$

Proof. Clearly, for each nonzero $q(z) \in \mathbb{F}[z]$, the map π_q is a projection map. Equation (5.3) follows by a simple computation. Since π_q is a projection, so is $I - \pi_q$. The identity $I = \pi_q + (I - \pi_q)$, taken together with (5.3), implies the direct sum (5.4). Finally, the isomorphism (5.6) follows from (5.5) and (5.3). ■

The isomorphism (5.6) allows us to lift the $\mathbb{F}[z]$-module structure to X_q. This module structure is the one induced by the polynomial z.

Definition 5.2. Let $q(z)$ be a monic polynomial in $\mathbb{F}[z]$. We define a linear transformation $S_q : X_q \longrightarrow X_q$ by

$$S_q f = z \cdot f = \pi_q z f. \tag{5.7}$$

We call S_q the **shift operator** in X_q. We note that for the shift operator S_q, we have for all $k \geq 0$, that $S_q^k f = \pi_q z^k f$. This implies that for any $p(z) \in \mathbb{F}[z]$, we have

$$p \cdot f = p(S_q) f = \pi_q(pf), \qquad f(z) \in X_q. \tag{5.8}$$

We refer to X_q, with the $\mathbb{F}[z]$-module structure induced by S_q, as a **polynomial model**.

The next proposition characterizes elements of a polynomial model and studies some important bases for it.

Proposition 5.3. *Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$. Then*

1. *A polynomial $f(z)$ belongs to X_q if and only if $q(z)^{-1}f(z)$ is strictly proper.*
2. *We have $\dim X_q = \deg q = n$.*
3. *The following sets are bases for X_q:*

   a. *The* **standard basis**, *namely $\mathscr{B}_{st} = \{1, z, \ldots, z^{n-1}\}$.*

b. The **control basis**, *namely* $\mathscr{B}_{co} = \{e_1(z), \ldots, e_n(z)\}$, *where*

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

c. In case $\alpha_1, \ldots, \alpha_n$, *the zeros of* $q(z)$, *are distinct then the polynomials* $p_i(z) = \Pi_{j \neq i}(z - \alpha_j)$, $i = 1, \ldots, n$ *form a basis for* X_q. *We refer to this as the* **spectral basis** $\mathscr{B}_{sp}$ *of* X_q.

d. Under the same assumption, the Lagrange interpolation polynomials, given by $l_i(z) = \frac{p_i(z)}{p_i(\alpha_i)}$ *are a basis for* X_q, *naturally called the* **interpolation basis**.

Proof. 1. Clearly, $f(z) \in X_q$ is equivalent to $\pi_q f = f$. In turn, this can be rewritten as $q^{-1}f = \pi_- q^{-1}f$, and this equality holds if and only if $q^{-1}f$ is strictly proper.

2. Clearly, the elements of X_q are all polynomials of degree $< n = \deg q$. Obviously, this is an n-dimensional space.

3. Each of the sets has n linearly independent elements, hence is a basis for X_q. The linear independence of the Lagrange interpolation polynomials was proved in Chapter 2, and the polynomials $p_i(z)$ are, up to a multiplicative constant, equal to the Lagrange interpolation polynomials. ■

We proceed by studying the matrix representations of S_q with respect to these bases of X_q.

Proposition 5.4. *Let* $S_q : X_q \longrightarrow X_q$ *be defined by (5.7).*

1. With respect to the standard basis, S_q *has the matrix representation*

$$C_q^{\sharp} = [S_q]_{st}^{st} = \begin{pmatrix} 0 & & -q_0 \\ 1 & & \cdot \\ & \cdot & \cdot \\ & & \cdot \\ & 1 & -q_{n-1} \end{pmatrix}. \tag{5.9}$$

2. With respect to the control basis, S_q *has the diagonal matrix representation*

$$C_q^{\flat} = [S_q]_{co}^{co} = \begin{pmatrix} 0 & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & \\ & & & 1 \\ -q_0 & \cdot & \cdot & -q_{n-1} \end{pmatrix}. \tag{5.10}$$

3. With respect to the spectral basis, S_q *has the matrix representation*

$$[S_q]_{sp}^{sp} = \begin{pmatrix} \alpha_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \alpha_n \end{pmatrix}. \tag{5.11}$$

Proof. 1. Clearly, we have

$$S_q z^i = \begin{cases} z^{i+1}, & i = 0, \dots, n-2, \\ -\sum_{i=0}^{n-1} q_i z^i, & i = n-1. \end{cases}$$

2. We compute, defining $e_0(z) = 0$,

$$\begin{aligned} S_q e_i &= \pi_q z e_i(z) = \pi_q z(z^{n-i} + q_{n-1} z^{n-i-1} + \cdots + q_i) \\ &= \pi_q(z^{n-i+1} + q_{n-1} z^{n-i} + \cdots + q_i z) \\ &= \pi_q(z^{n-i+1} + q_{n-1} z^{n-i} + \cdots + q_i z + q_{i-1}) - q_{i-1} e_n(z). \end{aligned}$$

So we get

$$S_q e_i = e_{i-1} - q_{i-1} e_n. \tag{5.12}$$

3. Noting that $q(z) = (z - \alpha_i) p_i(z)$, we compute

$$S_q p_i(z) = \pi_q(z - \alpha_i + \alpha_i) p_i = \pi_q(q + \alpha_i p_i) = \alpha_i p_i.$$

■

The matrices $C_q^\sharp, C_q^\flat$ are called the **companion matrices** of the polynomial $q(z)$.

We note that the change of basis transformation from the control to the standard basis has a particularly nice form. In fact, we have

$$[I]_{co}^{st} = \begin{pmatrix} q_1 & \cdot & \cdot & q_{n-1} & 1 \\ \cdot & & & \cdot & \\ \cdot & & \cdot & & \\ q_{n-1} & \cdot & & & \\ 1 & & & & \end{pmatrix}.$$

We relate now the invariant subspaces of the shift operators S_q, or equivalently the submodules of X_q, to factorizations of the polynomial $q(z)$. This is an example of the interplay between algebra and geometry that is one of the salient characteristics of the use of functional models.

Theorem 5.5. *Given a monic polynomial $q(z)$, $M \subset X_q$ is an S_q-invariant subspace if and only if*

$$M = q_1 X_{q_2},$$

for some factorization

$$q(z) = q_1(z) q_2(z).$$

Proof. Assume $q(z) = q_1(z)q_2(z)$ and $M = q_1 X_{q_2}$. Thus $f(z) \in M$ implies $f(z) = q_1(z)f_1(z)$ with $\deg f_1 < \deg q_2$. Using Lemma 1.22, we compute

$$S_q f = \pi_q z f = \pi_{q_1 q_2} z q_1 f_1 = q_1 \pi_{q_2} z f_1 = q_1 S_{q_2} f_1 \in M.$$

Conversely, let M be an S_q-invariant subspace. Now, for each $f(z) \in X_q$ there exists a scalar α that depends on $f(z)$ for which

$$S_q f = zf - \alpha q. \tag{5.13}$$

Consider now the set $N = M + q\mathbb{F}[z]$. Obviously N is closed under addition, and using (5.13), $z\{M + q\mathbb{F}[z]\} \subset \{M + q\mathbb{F}[z]\}$. Thus N is an ideal in $\mathbb{F}[z]$, hence of the form $q_1\mathbb{F}[z]$. Since, obviously, $q\mathbb{F}[z] \subset q_1\mathbb{F}[z]$, it follows from Proposition 1.46 that $q_1(z)$ is a divisor of $q(z)$, that is, we have a factorization $q(z) = q_1(z)q_2(z)$. It is clear that

$$M = \pi_q\{M + q\mathbb{F}[z]\} = \pi_q q_1 \mathbb{F}[z] = \pi_{q_1 q_2} q_1 \mathbb{F}[z] = q_1 X_{q_2}. \qquad \blacksquare$$

Proper subspaces of a vector space may have many complementary subspaces and invariant subspaces of polynomial models are no exception. The following proposition exhibits a particular complementary subspace to such a proper invariant subspace.

Proposition 5.6. *Let $q(z) \in \mathbb{F}[z]$ be nonzero and let $q(z) = q_1(z)q_2(z)$ be a factorization. Then, as vector spaces, we have the following direct sum representation:*

$$X_q = X_{q_1} \oplus q_1 X_{q_2}. \tag{5.14}$$

Proof. That $X_{q_1} \subset X_q$ follows from the fact that $\deg q_1 \leq \deg q$. Next, we note that $X_{q_1} \cap q_1 X_{q_2} = \{0\}$ and every nonzero polynomial in X_{q_1} has degree $< \deg q_1$, whereas a nonzero polynomial in $q_1 X_{q_2}$ has degree $\geq \deg q_1$. Finally, the equality in (5.14) follows from

$$\dim X_q = \deg q = \deg q_1 + \deg q_2 = \dim X_{q_1} + \dim X_{q_2} = \dim X_{q_1} \oplus q_1 X_{q_2}.$$

$\blacksquare$

Note that $X_{q_1} \subset X_q$ is generally not an invariant subspace for S_q.

The following proposition sums up the basic arithmetic properties of invariant subspaces of the shift operator. This can be viewed as the counterpart of Proposition 1.46.

Proposition 5.7. *Given a monic polynomial $q(z) \in \mathbb{F}[z]$, the following hold:*

1. *Let $q(z) = q_1(z)q_2(z) = p_1(z)p_2(z)$ be two factorizations. Then we have the inclusion*
$$q_1 X_{q_2} \subset p_1 X_{p_2} \tag{5.15}$$
if and only if $p_1(z) \mid q_1(z)$, or equivalently $q_2(z) \mid p_2(z)$.

2. *Given factorizations* $q(z) = p_i(z)q_i(z)$, $i = 1,\dots,s$, *then* $\cap_{i=1}^{s} p_i X_{q_i} = pX_q$ *with* $p(z)$ *the l.c.m of the* $p_i(z)$ *and* $q(z)$ *the g.c.d. of the* $q_i(z)$.
3. *Given factorizations* $q(z) = p_i(z)q_i(z)$, $i = 1,\dots,s$, *then* $\sum_{i=1}^{s} p_i X_{q_i} = pX_q$ *with* $q(z)$ *the l.c.m of the* $q_i(z)$ *and* $p(z)$ *the g.c.d. of the* $p_i(z)$.

Proof. 1. Assume $p_1(z)|q_1(z)$, i.e., $q_1(z) = p_1(z)r(z)$ for some polynomial $r(z)$. Then $q(z) = q_1(z)q_2(z) = (p_1(z)r(z))q_2(z) = p_1(z)(r(z)q_2(z)) = p_1(z)p_2(z)$, and in particular, $p_2(z) = r(z)q_2(z)$. This implies $q_1 X_{q_2} = p_1 r X_{q_2} \subset p_1 X_{rq_2} = p_1 X_{p_2}$.

Conversely, assume the inclusion (5.15) holds. From this we have

$$q_1 X_{q_2} + q\mathbb{F}[z] = q_1 X_{q_2} + q_1 q_2 \mathbb{F}[z] = q_1[X_{q_2} + q_2\mathbb{F}[z]] = q_1\mathbb{F}[z].$$

So (5.15) implies the inclusion $q_1\mathbb{F}[z] \subset p_1\mathbb{F}[z]$. By Proposition 1.46 it follows that $p_1(z)|q_1(z)$.

2. Let $t(z) = p_i(z)q_i(z)$, $i = 1,\dots,s$. Since $\cap_{i=1}^{s} p_i X_{q_i}$ is a submodule of X_q, it is of the form pX_q with $t(z) = p(z)q(z)$. Now the inclusion $pX_q \subset p_i X_{q_i}$ implies $p_i(z)|p(z)$ and $q(z)|q_i(z)$, so $p(z)$ is a common multiple of the $p_i(z)$, and $q(z)$ a common divisor of the $q_i(z)$. Let now $q'(z)$ be any common divisor of the $q_i(z)$. Since necessarily $q'(z) \mid t(z)$, we can write $t(z) = p'(z)q'(z)$. Now applying Proposition 1.46, we have $p'X_{q'} \subset p_i X_{q_i}$ and hence $p'X_{q'} \subset \cap_{i=1}^{s} p_i X_{q_i} = pX_q$. This implies $q'(z) \mid q(z)$, and hence $q(z)$ is a g.c.d. of the $q_i(z)$. By the same token, we conclude that $p(z)$ is the l.c.m. of the $p_i(z)$.
3. Since $p_1 X_{q_1} + \cdots + p_s X_{q_s}$ is an invariant subspace of X_t it is of the form pX_q with $t(z) = p(z)q(z)$. Now the inclusions $p_i X_{q_i} \subset pX_q$ imply the division relations $q_i(z) \mid q(z)$ and $p(z) \mid p_i(z)$. So $q(z)$ is a common multiple of the $q_i(z)$, and $p(z)$ a common divisor of the $p_i(z)$. Let $p'(z)$ be any other common divisor of the $p_i(z)$. Then $p_i(z) = p'(z)e_i(z)$ for some polynomials $e_i(z)$. Now $t(z) = p_i(z)q_i(z) = p'(z)e_i(z)q_i(z) = p'(z)q'(z)$, so $e_i(z)q_i(z) = q'(z)$, and $q'(z)$ is a common multiple of the $q_i(z)$. Now $p_i(z) = p'(z)e_i(z)$ implies $p_i X_{q_i} = p'e_i X_{q_i} \subset p'X_{e_i q_i} = p'X_{q'}$ and hence $pX_q = p_1 X_{q_1} + \cdots + p_s X_{q_s} = p'X_{q'}$. This shows that $q(z) \mid q'(z)$, and so $q(z)$ is the l.c.m. of the $q_i(z)$. ∎

Corollary 5.8. *Given the factorizations* $q(z) = p_i(z)q_i(z)$, $i = 1,\dots,s$, *then*

1.
$$X_q = p_1 X_{q_1} + \cdots + p_s X_{q_s}$$
if and only if the $p_i(z)$ *are coprime.*
2. *The sum* $p_1 X_{q_1} + \cdots + p_s X_{q_s}$ *is a direct sum if and only if* $q_1(z),\dots,q_s(z)$ *are mutually coprime.*
3. *We have the direct sum decomposition*
$$X_q = p_1 X_{q_1} \oplus \cdots \oplus p_s X_{q_s}$$
if and only if the $p_i(z)$ *are coprime and the* $q_i(z)$ *are mutually coprime.*

4. We have the direct sum decomposition

$$X_q = p_1 X_{q_1} \oplus \cdots \oplus p_s X_{q_s}$$

if and only if the $q_i(z)$ are mutually coprime and $q(z) = q_1(z)\cdots q_s(z)$. In this case $p_i(z) = \Pi_{j\neq i} q_j(z)$.

Proof. 1. Let the invariant subspace $p_1 X_{q_1} + \cdots + p_s X_{q_s}$ have the representation $p_\nu X_{q_\nu}$ with $p_\nu(z)$ the g.c.d. of the $p_i(z)$ and $q_\nu(z)$ the l.c.m. of the $q_i(z)$. Therefore $p_\nu X_{q_\nu} = X_q$ if and only if $p_\nu(z) = 1$ or equivalently $q_\nu(z) = q(z)$.

2. The sum $p_1 X_{q_1} + \cdots + p_s X_{q_s}$ is a direct sum if and only if for each index i, we have

$$p_i X_{q_i} \cap \sum_{j\neq i} p_j X_{q_j} = \{0\}.$$

Now $\sum_{j\neq i} p_j X_{q_j}$ is an invariant subspace and hence of the form $\pi_i X_{\sigma_i}$ for some factorization $q(z) = \pi_i(z)\sigma_i(z)$. Here $\pi_i(z)$ is the g.c.d. of the $p_j(z), j \neq i$ and $\sigma_i(z)$ is the l.c.m. of the $q_j(z)$, $j \neq i$. Now $p_i X_{q_i} \cap \pi_i X_{\sigma_i} = \{0\}$ if and only if $\sigma_i(z)$ and $q_i(z)$ are coprime. This, however, is equivalent to $q_i(z)$ being coprime with each of the $q_j(z)$, $j \neq i$, i.e., to the mutual coprimeness of the $q_i(z)$, $i = 1,\ldots,s$.

3. Follows from the previous two parts.

4. Clearly, the $p_i(z)$ are coprime. ■

Corollary 5.9. *Let $p(z) = p_1(z)^{\nu_1}\cdots p_k(z)^{\nu_k}$ be the primary decomposition of the polynomial $p(z)$. Define $\pi_i(z) = \Pi_{j\neq i} p_j(z)^{\nu_j}$, for $i = 1,\ldots,k$. Then*

$$X_p = \pi_1 X_{p_1^{\nu_1}} \oplus \cdots \oplus \pi_k X_{p_k^{\nu_k}}. \tag{5.16}$$

Proof. Clearly, the g.c.d. of the $\pi_i(z)$ is 1, whereas the l.c.m. of the $p_i^{\nu_i}(z)$ is $p(z)$. ■

The structure of the shift operator restricted to an invariant subspace can be easily deduced from the corresponding factorization.

Proposition 5.10. *Let $q(z) = q_1(z)q_2(z)$. Then we have the similarity*

$$S_q | q_1 X_{q_2} \simeq S_{q_2}. \tag{5.17}$$

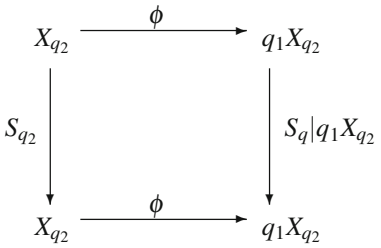
Proof. Let $\phi : X_{q_2} \longrightarrow q_1 X_{q_2}$ be the map defined by

$$\phi(f) = q_1 f,$$

which is clearly an isomorphism of the two spaces. Next we compute, for $f(z) \in X_{q_2}$,

$$\phi S_{q_2} f = q_1 S_{q_2} f = q\pi_q q_1 z f = \pi_q z q_1 f = S_q \phi f.$$

Therefore, the following diagram is commutative:



This is equivalent to (5.17). ■

Since eigenvectors span 1-dimensional invariant subspaces, we expect a characterization of eigenvectors of the shift S_q in terms of the polynomial $q(z)$.

Proposition 5.11. *Let $q(z)$ be a nonzero polynomial.*

1. *The eigenvalues of S_q coincide with the zeros of $q(z)$.*
2. *$f(z) \in X_q$ is an eigenvector of S_q corresponding to the eigenvalue α if and only if it has the representation*

$$f(z) = \frac{cq(z)}{z-\alpha}. \tag{5.18}$$

Proof. Let $f(z)$ be an eigenvector of S_q corresponding to the eigenvalue α, i.e., $S_q f = \alpha f$. By (5.13), there exists a scalar c for which $(S_q f)(z) = zf(z) - cq(z)$. Thus $zf(z) - cq(z) = \alpha f(z)$, which implies (5.18). Since $f(z)$ is a polynomial, we must have $q(\alpha) = 0$.

Conversely, if $q(\alpha) = 0$, $q(z)$ is divisible by $z-\alpha$ and hence $f(z)$, defined by (5.18), is in X_q. We compute

$$(S_q - \alpha I)f = \pi_q(z-\alpha)\frac{cq(z)}{z-\alpha} = \pi_q cq(z) = 0,$$

which shows that α is an eigenvalue of S_q, and a corresponding eigenvector is given by (5.18). ■

The previous proposition suggests that the characteristic polynomial of S_q is $q(z)$ itself. Indeed, this is true and is proved next.

Proposition 5.12. *Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$ be a monic polynomial of degree n, and let S_q be the shift operator defined by (5.7). Then the characteristic polynomial of S_q is $q(z)$.*

Proof. It suffices to compute $\det(zI - C)$ for an arbitrary matrix representation of S_q. We find it convenient to do the computation in the standard basis. In this case the matrix representation is given by the companion matrix $C_q^\sharp$ of (5.9). We prove the result by induction on n. For $n = 1$, we have $C = (-q_0)$ and $\det(zI - C) = z + q_0$.

Assume the statement holds up to $n-1$. We proceed to compute, using the induction hypothesis and expanding the determinant by the first row,

$$\det(zI - C_q^\sharp) = \begin{vmatrix} z & & & q_0 \\ -1 & & & \cdot \\ & . & & \cdot \\ & & . & \cdot \\ & & -1 & z+q_{n-1} \end{vmatrix}$$

$$= z \begin{vmatrix} z & & & q_1 \\ -1 & & & \cdot \\ & . & & \cdot \\ & & . & \cdot \\ & & -1 & z+q_{n-1} \end{vmatrix} + (-1)^{n+1} q_0 \begin{vmatrix} -1 & z & & \\ & . & . & \\ & & . & . \\ & & & -1 \end{vmatrix}$$

$$= z(z^{n-1} + q_{n-1}z^{n-2} + \cdots + q_1) + (-1)^{n+1} q_0 (-1)^{n-1}$$

$$= q(z). \qquad \blacksquare$$

We introduce now the class of cyclic transformations. Their importance lies in that they turn out to be the building blocks of general linear transformations.

Definition 5.13. Let $\mathscr{U}$ be an n-dimensional vector space over the field $\mathbb{F}$. A map $A : \mathscr{U} \longrightarrow \mathscr{U}$ is called a **cyclic transformation** if there exists a vector $b \in \mathscr{U}$ for which $\{b, Ab, \ldots, A^{n-1}b\}$ is a basis for $\mathscr{U}$. Such a vector b will be called a **cyclic vector** for A.

Lemma 5.14. *1. Given a nonzero polynomial $q(z)$ and $f(z) \in X_q$, the smallest S_q-invariant invariant subspace of X_q containing $f(z)$ is $q_1 X_{q_2}$, where $q_1(z) = q(z) \wedge f(z)$ and $q(z) = q_1(z)q_2(z)$.*

2. S_q is a cyclic transformation in X_q.

3. A polynomial $f(z) \in X_q$ is a cyclic vector of S_q if and only if $f(z)$ and $q(z)$ are coprime.

Proof. 1. Let M be the subspace of X_q spanned by the vectors $\{S_q^i f | i \geq 0\}$. This is the smallest S_q invariant subspace containing $f(z)$. Therefore it has the representation $M = q_1 X_{q_2}$ for a factorization $q(z) = q_1(z)q_2(z)$. Since $f(z) \in M$, there exists a polynomial $f_1(z) \in X_{q_2}$ for which $f(z) = q_1(z)f_1(z)$. This shows that $q_1(z)$ is a common divisor of $q(z)$ and $f(z)$.

To show that it is the greatest common divisor, let us assume that $q'(z)$ is an arbitrary common divisor of $q(z)$ and $f(z)$. Thus we have $q(z) = q'(z)q''(z)$ and $f(z) = q'(z)f'(z)$. Using Lemma 1.24, we compute

$$S_q^k f = \pi_q x^k f = \pi_q x^k q' f' = q' \pi_{q''} x^k f' = q' S_{q''}^k f'.$$

Thus we have $M \subset q' X_{q''}$, or equivalently $q_1 X_{q_2} \subset q' X_{q''}$. This implies $q'(z)|q_1(z)$; hence $q_1(z)$ is the g.c.d. of $q(z)$ and $f(z)$.
2. Obviously $1 \in X_q$ and $1 \wedge q(z) = 1$. So 1 is a cyclic vector for S_q.
3. Obviously, since $\dim q_1 X_{q_2} = \dim X_{q_2} = \deg q_2$, $X = q_1 X_{q_2}$ if and only if $\deg q_1 = 0$, that is, $f(z)$ and $q(z)$ are coprime. ■

The availability of eigenvectors of the shift allows us to study under what conditions the shift S_q is diagonalizable, i.e., has a diagonal matrix representation.

Proposition 5.15. *Let $q(z)$ be a monic polynomial of degree n. Then S_q is diagonalizable if and only if $q(z)$ splits into the product of n distinct linear factors, or equivalently, it has n distinct zeros.*

Proof. Assume $\alpha_1, \ldots, \alpha_n$ are the distinct zeros of $q(z)$, i.e., $q(z) = \Pi_{i=1}^n (z - \alpha_i)$. Let $p_i(z) = \frac{q(z)}{z-\alpha_i} = \Pi_{j \neq i}(z - \alpha_j)$. Then $\mathscr{B}_{sp} = \{p_1, \ldots, p_n\}$ is the spectral basis for X_q, differing from the Lagrange interpolation basis by constant factors only. It is easily checked that $(S_q - \alpha_i) p_i = 0$. So

$$[S_q]_{sp}^{sp} = \begin{pmatrix} \alpha_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \alpha_n \end{pmatrix}, \tag{5.19}$$

and S_q is diagonalizable.

Conversely, assume S_q is diagonalizable. Then with respect to some basis it has the representation (5.19). Since S_q is cyclic, its minimal and characteristic polynomials coincide. Necessarily all the α_i are distinct. ■

Proposition 5.16. *Let $q(z)$ be a monic polynomial and S_q the shift operator in X_q defined by (5.7). Then*

$$p(S_q) f = \pi_q(pf), \qquad f(z) \in X_q. \tag{5.20}$$

Proof. Using linearity, it suffices to show that

$$S_q^k f = \pi_q z^k f, \qquad f(z) \in X_q.$$

We prove this by induction. For $k = 1$ this is the definition. Assume we proved it up to an integer k. Now, using the fact that $z \mathrm{Ker}\, \pi_q \subset \mathrm{Ker}\, \pi_q$, we compute

$$S_q^{k+1} f = S_q S_q^k f = \pi_q z \pi_q z^k f = \pi_q z^{k+1} f.$$

■

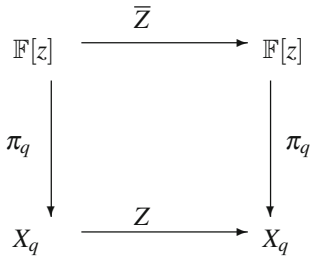
Clearly, the operators $p(S_q)$ all commute with the shift S_q. We proceed to state the simplest version of the commutant lifting theorem. It characterizes operators

commuting with the shift S_q in X_q via operators commuting with the shift S_+ in $\mathbb{F}[z]$. The last class of operators consists of multiplication operators by polynomials.

Theorem 5.17. *1. Let $q(z)$ be a monic polynomial and S_q the shift operator in X_q defined by (5.7). Let Z be any operator in X_q that commutes with S_q. Then there exists an operator $\overline{Z}$ that commutes with S_+ and such that*

$$Z = \pi_q \overline{Z}|_{X_q}. \tag{5.21}$$

Equivalently, the following diagram is commutative:



2. An operator Z in X_q commutes with S_q if and only if there exists a polynomial $p(z) \in \mathbb{F}[z]$ for which

$$Zf = \pi_q pf = p(S_q)f, \qquad f(z) \in X_q. \tag{5.22}$$

Proof. 1. By Proposition 6.2, there exists a polynomial $p(z)$ for which $Z = p(S_q)$. We define $\overline{Z} : \mathbb{F}[z] \longrightarrow \mathbb{F}[z]$ by $\overline{Z}f = pf$. It is easily checked that (5.21) holds.

2. If Z is represented by (5.22), then

$$S_q Zf = \pi_q z \pi_q pf = \pi_q zpf = \pi_q pzf = \pi_q p \pi_q zf = ZS_q f.$$

Conversely, if Z commutes with the shift, it has a representation (5.21). Now, any map $\overline{Z}$ commuting with S_+ is a multiplication by a polynomial $p(z)$, which proves (5.22). ■

The proof of the previous theorem is slightly misleading, inasmuch as it uses the fact that S_q is cyclic. In Chapter 8, we will return to this subject using tensor products and the Bezout map.

Since any operator of the form $p(S_q)$ is completely determined by the two polynomials $p(z)$ and $q(z)$, all its properties should be derivable from these data. The next proposition focuses on the invertibility properties of $p(S_q)$.

Theorem 5.18. *Given the polynomials $p(z), q(z)$, with $q(z)$ monic, let $r(z)$ and $s(z)$ be the g.c.d and the l.c.m. of $p(z)$ and $q(z)$ respectively. Then*

1. *We have the factorizations*

$$\begin{aligned} q(z) &= r(z)q_1(z), \\ p(z) &= r(z)p_1(z), \end{aligned} \tag{5.23}$$

with $p_1(z), q_1(z)$ coprime, as well as the factorizations

$$s(z) = r(z)p_1(z)q_1(z) = p(z)q_1(z) = q(z)p_1(z). \tag{5.24}$$

Moreover, we have

$$\begin{aligned} \operatorname{Ker} p(S_q) &= q_1 X_r, \\ \operatorname{Im} p(S_q) &= r X_{q_1}. \end{aligned} \tag{5.25}$$

2. *Let $p(z), q(z) \in \mathbb{F}[z]$, with $q(z)$ nonzero. Then the linear transformation $p(S_q)$ is invertible if and only if $p(z)$ and $q(z)$ are coprime. Moreover, we have*

$$p(S_q)^{-1} = a(S_q), \tag{5.26}$$

where the polynomial $a(z)$ arises out of any solution of the Bezout equation

$$a(z)p(z) + b(z)q(z) = 1. \tag{5.27}$$

Proof. 1. With $r(z)$ and $s(z)$ defined as above, (5.24) is an immediate consequence of (5.23). Applying Theorem 5.5, it follows that $q_1 X_r$ and $r X_{q_1}$ are invariant subspaces of X_q. If $f(z) \in q_1 X_r$,, then $f = q_1 g$, with $g \in X_r$. We compute, using (5.23),

$$p(S_q)f = \pi_q p f = q_1 r \pi_- r^{-1} q_1^{-1} r p_1 q_1 g = q_1 r \pi_- g = 0,$$

which shows that $q_1 X_r \subset \operatorname{Ker} p(S_q)$.

Conversely, assume $f(z) \in \operatorname{Ker} p(S_q)$. Then $\pi_q p f = 0$ or there exists $g(z)$ such that $p(z)f(z) = q(z)g(z)$, which implies $p_1(z)f(z) = q_1(z)g(z)$. Since $p_1(z)$ and $q_1(z)$ are coprime, we have $f(z) = q_1(z)f_1(z)$ for some polynomial $f_1(z)$. As $f(z) \in X_q$, we must have $f_1(z) \in X_r$. So $\operatorname{Ker} p(S_q) \subset q_1 X_r$, and the first equality in (5.25) follows.

Next, assume $g(z) \in \operatorname{Im} p(S_q)$, i.e., there exists an $f(z) \in X_q$ such that $g = \pi_q p f$. We compute, using (5.23) and (5.24),

$$g = \pi_q p f = q_1 r \pi_- r^{-1} q_1^{-1} r p_1 f = r \pi_{q_1} p_1 f \in r X_{q_1},$$

i.e., we have the inclusion $\operatorname{Im} p(S_q) \subset r X_{q_1}$.

Conversely, assume $g(z) \in r X_{q_1}$, i.e., $g(z) = r(z)g_1(z)$ with $g_1(z) \in X_{q_1}$. By the coprimeness of $p_1(z)$ and $q_1(z)$, the map $f_1 \mapsto \pi_{q_1} p_1 f_1$, acting in X_{q_1}, is an invertible map. Hence, there exists $f_1(z) \in X_{q_1}$ for which $g_1 = \pi_{q_1} p_1 f_1$. This implies

$$rg_1 = rq_1\pi_- r^{-1}q_1^{-1}p_1 rf_1 = \pi_q prf_1.$$

This shows that $rX_{q_1} \subset \operatorname{Im} p(S_q)$, and the second equality in (5.25) follows.

2. From the characterization (5.25) of $\operatorname{Ker} p(S_q)$ and $\operatorname{Im} p(S_q)$, it follows that the injectivity of $p(S_q)$ is equivalent to the coprimeness of $p(z)$ and $q(z)$, and the same holds for surjectivity.

   In order to actually invert $p(S_q)$, we use the coprimeness of $p(z)$ and $q(z)$. This implies the existence of polynomials $a(z), b(z)$ that solve the Bezout equation $a(z)p(z) + b(z)q(z) = 1$. Applying the functional calculus to the shift S_q, and noting that $q(S_q) = 0$, we get

$$a(S_q)p(S_q) + b(S_q)q(S_q) = a(S_q)p(S_q) = I,$$

   and (5.26) follows. ■

5.3 Circulant Matrices

In this section we give a short account of a special class of structured matrices, called circulant matrices. We do this for its own sake, and as an illustration of the power of polynomial algebra.

Definition 5.19. An $n \times n$ matrix over a field $\mathbb{F}$ is called a **circulant matrix**, or simply **circulant** for short, if it has the form

$$C = \operatorname{circ}(c_0, \ldots, c_{n-1}) = \begin{pmatrix} c_0 & c_{n-1} & . & . & c_1 \\ c_1 & c_0 & . & . & . \\ . & . & . & . & . \\ . & . & . & . & c_{n-1} \\ c_{n-1} & . & . & c_1 & c_0 \end{pmatrix},$$

i.e., $c_{ij} = c_{(i-j \bmod n)}$. We define a polynomial $c(z)$ by $c(z) = c_0 + c_1 z + \cdots + c_{n-1}z^{n-1}$. The polynomial c is called the **representer** of $\operatorname{circ}(c_0, \ldots, c_{n-1})$.

Theorem 5.20. *For circulant matrices the following properties hold:*

1. *The circulant matrix* $\operatorname{circ}(c_0, \ldots, c_{n-1})$ *is the matrix representation of* $c(S_{z^n-1})$ *with respect to the standard basis of* X_{z^n-1}.
2. *The sum of circulant matrices is a circulant matrix. Specifically,*

$$\operatorname{circ}(a_1, \ldots, a_n) + \operatorname{circ}(b_1, \ldots, b_n) = \operatorname{circ}(a_1 + b_1, \ldots, a_n + b_n).$$

3. *We have*

$$\alpha \operatorname{circ}(a_1, \ldots, a_n) = \operatorname{circ}(\alpha a_1, \ldots, \alpha a_n).$$

4. *Define the special circulant matrix Π by*

$$\Pi = \operatorname{circ}(0,1,0,\ldots,0) = \begin{pmatrix} 0 & & & & 1 \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}.$$

   Then $C \in \mathbb{F}^{n\times n}$ is a circulant if and only if $C\Pi = \Pi C$.
5. *The product of circulants is commutative.*
6. *The product of circulants is a circulant.*
7. *The inverse of a circulant is a circulant. Moreover, the inverse of a circulant with representer $c(z)$ is a circulant with representer $a(z)$, where $a(z)$ comes from a solution of the Bezout equation $a(z)c(z)+b(z)(z^n-1)=1$.*
8. *Over an algebraically closed field, circulants are diagonalizable.*

Proof. 1. We compute

$$[S_{z^n-1}]_{st}^{st} = \begin{pmatrix} 0 & \ldots & 1 \\ 1 & & \cdot \\ & \ddots & \cdot \\ & & \cdot \\ & 1 & 0 \end{pmatrix}.$$

This is a special case of the companion matrix in (5.9). Now

$$S_{z^n-1}^i z^j = \begin{cases} z^{i+j}, & i+j \le n-1, \\ z^{i+j-n}, & i+j \ge n. \end{cases}$$

So

$$c(S_{z^n-1})z^j = \sum_{i=0}^{n-1} c_i S_{z^n-1}^i z^j = \sum_{i=0}^{n-1-j} c_i z^{i+j} + \sum_{i=n-j}^{n-1} c_i z^{i+j-n},$$

and this implies the equality

$$[c(S_{z^n-1})]_{st}^{st} = \operatorname{circ}(c_0,\ldots,c_{n-1}).$$

2. Given polynomials $a(z), b(z) \in \mathbb{F}[z]$, we have

$$(a+b)(S_{z^n-1}) = a(S_{z^n-1}) + b(S_{z^n-1}),$$

and hence

$$\begin{aligned} \operatorname{circ}(a_0+b_0,\ldots,a_{n-1}+b_{n-1}) &= [(a+b)(S_{z^n-1})]_{st}^{st} \\ &= [a(S_{z^n-1})]_{st}^{st} + [b(S_{z^n-1})]_{st}^{st} \\ &= \operatorname{circ}(a_0,\ldots,a_{n-1}) + \operatorname{circ}(b_0,\ldots,b_{n-1}). \end{aligned}$$

3. We compute

$$\operatorname{circ}(\alpha a_0,\dots,\alpha a_{n-1}) = [\alpha a(S_{z^n-1})]_{st}^{st} = \alpha[a(S_{z^n-1})]_{st}^{st} = \alpha\operatorname{circ}(a_0,\dots,a_{n-1}).$$

4. Clearly, $\Pi = \operatorname{circ}(0,1,0,\dots,0) = [S_{z^n-1}]_{st}^{st}$. Obviously, S_{z^n-1} is cyclic. Hence, a linear transformation K commutes with S_{z^n-1} if and only if $K = c(S_{z^n-1})$ for some polynomial $c(z)$. Thus, assume $C\Pi = \Pi C$. Then there exists a linear transformation K in X_{z^n-1} satisfying $[K]_{st}^{st} = C$, and K commutes with S_{z^n-1}. Therefore $K = c(S_{z^n-1})$ and $C = \operatorname{circ}(c_0,\dots,c_{n-1})$.
   Conversely, if $C = \operatorname{circ}(c_0,\dots,c_{n-1})$, we have

$$C\Pi = [c(S_{z^n-1})]_{st}^{st}[S_{z^n-1}]_{st}^{st} = [S_{z^n-1}]_{st}^{st}[c(S_{z^n-1})]_{st}^{st} = \Pi C.$$

5. Follows from

$$c(S_{z^n-1})d(S_{z^n-1}) = d(S_{z^n-1})c(S_{z^n-1}).$$

6. Follows from

$$c(S_{z^n-1})d(S_{z^n-1}) = (cd)(S_{z^n-1}).$$

7. Let $C = \operatorname{circ}(c_0,\dots,c_{n-1}) = c(S_{z^n-1})$, where $c(z) = c_0 + c_1 z + \cdots + c_{n-1}z^{n-1}$. By Theorem 5.18, $c(S_{z^n-1})$ is invertible if and only if $c(z)$ and $z^n - 1$ are coprime. In this case there exist polynomials $a(z), b(z)$ satisfying the Bezout identity $a(z)c(z) + b(z)(z^n - 1) = 1$. We may assume without loss of generality that $\deg a < n$. From the Bezout identity we conclude that $c(S_{z^n-1})c(S_{z^n-1}) = I$ and hence

$$\operatorname{circ}(c_0,\dots,c_{n-1})^{-1} = [a(S_{z^n-1})]_{st}^{st} = \operatorname{circ}(a_0,\dots,a_{n-1}).$$

8. The polynomial $z^n - 1$ has a multiple zero if and only if $z^n - 1$ and nz^{n-1} have a common zero. Clearly, this cannot occur. Since all the roots of $z^n - 1$ are distinct, it follows from Proposition 5.15 that S_{z^n-1} is diagonalizable. This implies the diagonalizability of $c(S_{z^n-1})$. ■

5.4 Rational Models

Given a field $\mathbb{F}$, we saw that the ring of polynomials $\mathbb{F}[z]$ is an entire ring. Hence, by Theorem 1.49, it is embeddable in its field of quotients. We call the field of quotients of $\mathbb{F}[z]$ the field of **rational functions** and denote it by $\mathbb{F}(z)$. Strictly speaking, the elements of $\mathbb{F}(z)$ are equivalence classes of pairs of polynomials $(p(z), q(z))$, with $q(z)$ nonzero. However, in each nonzero equivalence class, there is a unique pair with $p(z), q(z)$ coprime and $q(z)$ monic. The corresponding equivalence class will be denoted by $\frac{p(z)}{q(z)}$. Given such a pair of polynomials $p(z), q(z)$, there is a unique representation of $p(z)$ in the form $p(z) = a(z)q(z) + r(z)$, with $\deg r < \deg q$. This allows us to write

$$\frac{p(z)}{q(z)} = a(z) + \frac{r(z)}{q(z)}. \tag{5.28}$$

A rational function $r(z)/q(z)$ with $\deg r \leq \deg q$ will be called **proper**, and if $\deg r < \deg q$ is satisfied, **strictly proper**. Thus, any rational function $g(z) = \frac{p(z)}{q(z)}$ has a unique representation as a sum of a polynomial and a strictly proper rational function. We denote by $\mathbb{F}_-(z)$ the space of strictly proper rational functions and observe that it is an infinite-dimensional linear space. Equation (5.28) means that for $\mathbb{F}(z)$ we have the following direct sum decomposition:

$$\mathbb{F}(z) = \mathbb{F}[z] \oplus \mathbb{F}_-(z). \tag{5.29}$$

With this direct sum decomposition we associate two projection operators in $\mathbb{F}(z)$, π_+ and π_-, with images $\mathbb{F}[z]$ and $\mathbb{F}_-(z)$ respectively. To be precise, given the representation (5.28), we have

$$\begin{aligned} \pi_+\left(\frac{p}{q}\right) &= a \\ \pi_-\left(\frac{p}{q}\right) &= \frac{r}{q}. \end{aligned} \tag{5.30}$$

Proper rational functions have an expansion as formal power series in the variable z^{-1}, i.e., in the form $g(z) = \sum_{i=0}^{\infty} \frac{g_i}{z^i}$. Assume $p(z) = \sum_{k=0}^{n} p_k z^k, q(z) = \sum_{i=0}^{n} q_i z^i$, with $q_n \neq 0$. We compute

$$\sum_{k=0}^{n} p_k z^k = \sum_{i=0}^{n} q_i z^i \sum_{j=0}^{\infty} \frac{g_j}{z^j} = \sum_{k=-\infty}^{n} \left\{ \sum_{i=k}^{n} q_i g_{i-k} \right\} z^k.$$

By comparing coefficients we get the infinite system of linear equations

$$p_k = \sum_{i=k}^{n} q_i g_{i-k}, \qquad -\infty < k \leq n.$$

This system has a unique solution, which can be found by solving it recursively, starting with $k = n$. An alternative way of finding this expansion is via the process of long division of $p(z)$ by $q(z)$.

This generalizes easily to the case of the field of rational functions. We can consider the field $\mathbb{F}(z)$ of rational functions as a subfield of $\mathbb{F}((z^{-1}))$, the field of truncated Laurent series. The space $\mathbb{F}_-(z)$ of strictly proper rational functions can be viewed as a subspace of $z^{-1}\mathbb{F}[[z^{-1}]]$. In the same way that we have the isomorphism $z^{-1}\mathbb{F}[[z^{-1}]] \simeq \mathbb{F}((z^{-1}))/\mathbb{F}[z]$, we have also $\mathbb{F}_-(z) \simeq \mathbb{F}(z)/\mathbb{F}[z]$. In fact, the projection $\pi_- : \mathbb{F}(z) \longrightarrow \mathbb{F}_-(z)$ is a surjective linear map with kernel equal to $\mathbb{F}[z]$. However, the spaces $\mathbb{F}((z^{-1}))$, $\mathbb{F}(z)$, and $\mathbb{F}[z]$ all carry also a natural $\mathbb{F}[z]$-module structure, with polynomials acting by multiplication. This structure induces an $\mathbb{F}[z]$-module structure in the quotient spaces. This module structure is transferred to $\mathbb{F}_-(z)$ by defining, for $p(z) \in \mathbb{F}[z]$,

$$p \cdot g = \pi_-(pg), \qquad g(z) \in \mathbb{F}_-(z). \tag{5.31}$$

In particular, we define the **backward shift operator** S_-, acting in $\mathbb{F}_-(z)$, by

$$S_- g = \pi_-(zg), \qquad g \in \mathbb{F}_-(z). \tag{5.32}$$

We shall use the same notation when $z^{-1}\mathbb{F}[[z^{-1}]]$ replaces $\mathbb{F}_-(z)$. We point out that in the behavioral literature, S_- is denoted by σ. Using this, equation (5.31) can be rewritten, for $p(z) \in \mathbb{F}[z]$, as a Toeplitz-like operator

$$p(\sigma)g = \pi_-(pg), \qquad g(z) \in z^{-1}\mathbb{F}[[z^{-1}]]. \tag{5.33}$$

In terms of the expansion of $g(z) \in \mathbb{F}_-(z)$ around infinity, i.e., in terms of the representation $g(z) = \sum_{i=1}^{\infty} \frac{g_i}{z^i}$ we have

$$S_- \sum_{i=1}^{\infty} \frac{g_i}{z^i} = \sum_{i=1}^{\infty} \frac{g_{i+1}}{z^i}.$$

This explains the use of the term backward shift for this operator.

We will show now that for the backward shift S_- in $z^{-1}\mathbb{F}[[z^{-1}]]$, any $\alpha \in \mathbb{F}$ is an eigenvalue and it has associated eigenfunctions of any order. Indeed, let $\alpha \in \mathbb{F}$. Then it is easy to check that $(z-\alpha)^{-1} = \sum_{i=1}^{\infty} \frac{\alpha^{i-1}}{z^i}$ is an eigenvector of S_- correspomding to the eigenvalue α. Analogously, $(z-\alpha)^{-k}$ is a generalized eigenvector of S_- of order k. Moreover, any other such eigenvector is necessarily of the form $c(z-\alpha)^{-k}$. Therefore, we have, for $k \geq 1$, $\dim \operatorname{Ker} (\alpha I - S_-)^k = k$.

In contrast to this richness of eigenfunctions of the backward shift S_-, the forward shift S_+ does not have any eigenfunctions. This asymmetry will be further explored in the study of duality.

The previous discussion indicates that the spectral structure of the shift S_-, with finite multiplicity, might be rich enough to model all finite-dimensional linear transformations, up to similarity transformations, as the backward shift S_- restricted to a finite-dimensional backward-shift-invariant subspace. This turns out to be the case, and this result lies at the heart of the matter, explaining the effectiveness of the use of shifts in linear algebra and system theory. In fact, the whole book can be considered, to a certain extent, to be an elaboration of this idea. We will study a special case of this in Chapter 6.

It is easy to construct many finite-dimensional S_--invariant subspaces of $\mathbb{F}_-(z)$. In fact, given any nonzero polynomial $d(z)$, we let

$$X^d = \left\{ \frac{r}{d} \mid \deg r < \deg d \right\}.$$

It is easily checked that X^d is indeed an S_--invariant subspace, and its dimension equals the degree of $d(z)$.

It is natural to consider the restriction of the operator S_- to X^d. Thus we define a linear transformation $S^d : X^d \longrightarrow X^d$ by $S^d = S_-|X^d$, or equivalently, for $h \in X^d$,

$$S^d h = S_- h = \pi_- z h. \tag{5.34}$$

The modules X_d and X^d have the same dimension and are defined by the same polynomial. Just as the polynomial model X_d has been defined in terms of the projection π_d, so can the rational model X^d be characterized as the image of a projection.

Definition 5.21. Let the map $\pi^d : z^{-1}\mathbb{F}[[z^{-1}]] \longrightarrow z^{-1}\mathbb{F}[[z^{-1}]]$ be defined by

$$\pi^d h = \pi_- d^{-1}\pi_+ dh, \qquad h(z) \in \mathbb{F}[[z^{-1}]]. \tag{5.35}$$

Proposition 5.22. *Let π^d be defined by (5.35). Then*

1. *π^d is a projection in $z^{-1}\mathbb{F}[[z^{-1}]]$.*
2. *We have*
$$X^d = \operatorname{Im}\pi^d. \tag{5.36}$$
3. *With $d(\sigma)$ defined by (5.33), we have*
$$X^d = \operatorname{Ker} d(\sigma). \tag{5.37}$$

Proof. 1. For $h(z) \in z^{-1}\mathbb{F}[[z^{-1}]]$, we compute

$$\begin{aligned}(\pi^d)^2 h &= \pi^d(\pi^d h) = \pi_- d^{-1}\pi_+ d\pi_- d^{-1}\pi_+ dh \\ &= \pi_- d^{-1}\pi_+ dd^{-1}\pi_+ dh = \pi_- d^{-1}\pi_+\pi_+ dh \\ &= \pi_- d^{-1}\pi_+ dh = \pi^d h,\end{aligned}$$

i.e., π^d is a projection.

2. Assume $h(z) = d^{-1}r$, with $\deg r < \deg d$. Then

$$\pi^d(d^{-1}r) = \pi_- d^{-1}\pi_+ dd^{-1}r = \pi_- d^{-1}r = d^{-1}r,$$

which shows that $X^d \subset \operatorname{Im}\pi^d$.

Conversely, assume $h(z) \in \operatorname{Im}\pi^d$. Thus there exists $g(z) \in z^{-1}\mathbb{F}[[z^{-1}]]$ for which $h(z) = \pi^d g$. Hence

$$\pi^d h = (\pi^d)^2 g = \pi^d g = h.$$

From this it follows that

$$\pi_- dh = \pi_- d\pi^d h = \pi_- d\pi_- d^{-1}\pi_+ dh = \pi_- dd^{-1}\pi_+ dh = \pi_-\pi_+ dh = 0.$$

This implies that there exists $r(z) \in \mathbb{F}[z]$ for which $d(z)h(z) = r(z)$ and $\deg r < \deg d$. In turn, we conclude that $h(z) = d(z)^{-1}r(z) \in X^d$, which implies the inclusion $\operatorname{Im}\pi^d \subset X^d$.

3. Assume $h \in \operatorname{Ker} d(\sigma)$, which implies $\pi_+ dh = dh$. We compute

$$\pi^d h = \pi_- d^{-1}\pi_+ dh = \pi_- d^{-1}dh = \pi_- h = h,$$

i.e., $\operatorname{Ker} d(\sigma) \subset X^d$.

Conversely, assume $h \in X^d$. This implies $h = \pi_- d^{-1}\pi_+ dh$. We compute

$$d(\sigma)h = \pi_- dh = \pi_- d\pi_- d^{-1}\pi_+ dh = \pi_- dd^{-1}\pi_+ dh = \pi_-\pi_+ dh = 0,$$

i.e., $X^d \subset \operatorname{Ker} d(\sigma)$. The two inclusions imply (5.37). ■

It is natural to conjecture that the polynomial model X_d and the rational model X^d must be isomorphic, and this is indeed the case, as the next theorem shows.

Theorem 5.23. *Let $d(z)$ be a nonzero polynomial. Then the operators S_d and S^d are isomorphic. The isomorphism is given by the map $\rho_d : X^d \longrightarrow X_d$ defined by*

$$\rho_d h = dh. \tag{5.38}$$

Proof. We compute

$$\rho_d S^d h = \rho_d \pi_- zh = d\pi_- zh = d\pi_- d^{-1}dzh = \pi_d z(dh) = S_d(\rho_d h),$$

i.e.,

$$\rho_d S^d = S_d \rho_d, \tag{5.39}$$

which, by the invertibility of ρ_d, proves the isomorphism. ■

The polynomial and rational models that have been introduced are isomorphic, yet they represent two fundamentally different points of view. In the case of polynomial models, all spaces of the form X_q, with $\deg q = n$, contain the same elements, but the associated shifts S_q act differently. On the other hand, the operators S^q in the spaces X^q act in the same way, since they are all restrictions of the backward shift S_-. However, the spaces are different. Thus the polynomial models represent an arithmetic perspective, whereas rational models represent a geometric one.

Our next result is the characterization of all finite-dimensional S_--invariant subspaces.

Proposition 5.24. *A subset M of $\mathbb{F}_-(z)$ is a finite-dimensional S_--invariant subspace if and only if we have $M = X^d$.*

Proof. Assume that for some polynomial $d(z)$, $M = X^d$. Then

$$S_-\frac{r}{d} = \pi_- z\frac{r}{d} = d^{-1}d\pi_- d^{-1}zr = d^{-1}\pi_d zr \in M.$$

Conversely, let M be a finite-dimensional S_--invariant subspace. By Theorem 4.54, there exists a nonzero polynomial $p(z)$ of minimal degree such that $\pi_- ph = 0$, for all $h \in M$. Thus, we get $M \subset X^p$. It follows that pM is a submodule of X_p, hence of the form $p_1 X_{p_2}$ for some factorization $p(z) = p_1(z)p_2(z)$. We conclude that $M = p^{-1}p_1 X_{p_2} = p_2^{-1}p_1^{-1}p_1 X_{p_2} = X^{p_2}$. The minimality of $p(z)$ implies $p(z) = p_2(z)$, up to a constant nonzero factor. ■

The spaces of rational functions of the form X^d will be referred to as **rational models**. Note that in the theory of differential equations, these spaces appear as the Laplace transforms of the spaces of solutions of a homogeneous linear differential equation with constant coefficients.

The following sums up the basic arithmetic properties of rational models. It is the counterpart of Proposition 5.7.

Proposition 5.25. *1. Given polynomials $p(z), q(z) \in \mathbb{F}[z]$, we have the inclusion $X^p \subset X^q$ if and only if $p(z) \mid q(z)$.*
2. Given polynomials $p_i(z) \in \mathbb{F}[z]$, $i = 1, \ldots, s$, then $\cap_{i=1}^s X^{p_i} = X^p$ with $p(z)$ the g.c.d. of the $p_i(z)$.
3. Given polynomials $p_i \in \mathbb{F}[z]$, then $\sum_{i=1}^s X^{p_i} = X^q$ with $q(z)$ the l.c.m. of the $p_i(z)$.

Proof. Follows from Proposition 5.7, using the isomorphism of polynomial and rational models given by Theorem 5.23. ■

The primary decomposition theorem and the direct sum representation (5.16) have a direct implication toward the partial fraction decomposition of rational functions.

Theorem 5.26. *Let $p(z) = \Pi_{i=1}^s p_i(z)^{\nu_i}$ be the primary decomposition of the nonzero polynomial $p(z)$. Then*

1. We have

$$X^p = X^{p_1^{\nu_1}} \oplus \cdots \oplus X^{p_s^{\nu_s}}. \tag{5.40}$$

2. Each rational function $g(z) \in X^p$ has a unique representation of the form

$$g(z) = \frac{r(z)}{p(z)} = \sum_{i=1}^{s}\sum_{j=1}^{\nu_i}\frac{r_{ij}}{p_i^j},$$

with $\deg r_{ij} < \deg p_i$, $j = 1, \ldots, \nu_i$.

Proof. 1. Given the primary decomposition of $p(z)$, we define $\pi_i(z) = \Pi_{j\neq i} p_i^{\nu_i}$. Clearly, by Corollary 5.9, we have

$$X_p = \pi_1 X_{p_1^{\nu_1}} \oplus \cdots \oplus \pi_s X_{p_s^{\nu_s}}. \tag{5.41}$$

We use the isomorphism of the modules X_p and X^p and the fact that $p(z) = \pi_i(z)p_i(z)^{\nu_i}$, which implies $p^{-1}\pi_i X_{p_i^{\nu_i}} = p_i^{-\nu_i}\pi_i - 1\pi_i X_{p_i^{\nu_i}} = X^{p_i^{\nu_i}}$, to get the direct sum decomposition (5.40).

2. For any $r_i(z) \in X_{p_i^{\nu_i}}$ we have $r_i = \sum_{j=0}^{\nu_i - 1} r_{i(\nu_i - j)} p_i^j$. With $\frac{r}{p} = \sum_{i=1}^{s} \frac{r_i}{p_i^{\nu_i}}$, using (5.41) and taking $r_i \in X_{p_i^{\nu_i}}$, we have

$$\frac{r_i}{p_i^{\nu_i}} = \sum_{j=1}^{\nu_i} \frac{r_{ij}}{p_i^j}. \tag{5.42}$$

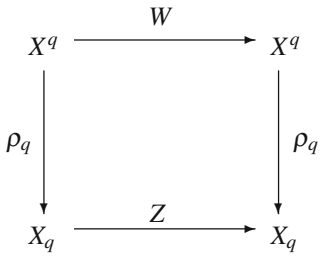
■

The isomorphism (5.39) between the shifts S_q and S^q can be used to transform Theorems 5.17 and 5.18 into the context of rational models. Thus we have the following Theorem.

Theorem 5.27. *1. Let $q(z)$ be a monic polynomial and S^q the shift operator in X^q defined by (5.34). Let W be any operator in X^q that commutes with S^q. Then there exists an operator Z that commutes with S_q and such that*

$$Z = \rho_q W \rho_q^{-1}. \tag{5.43}$$

Equivalently, the following diagram is commutative:



2. An operator W in X^q commutes with S^q if and only if there exists a polynomial $p(z) \in \mathbb{F}[z]$ for which

$$Wh = \pi_- ph = p(S^q)h, \qquad h(z) \in X^q. \tag{5.44}$$

Proof. 1. By the isomorphism (5.39), we have $\rho_q S^q = S_q \rho_q$. Since $WS^q = S^q W$, it follows that $W\rho_q^{-1} S_q \rho_q = \rho_q^{-1} S_q \rho_q W$, or $(\rho_q W \rho_q^{-1})S_q = S_q(\rho_q W \rho_q^{-1})$. Defining $Z = \rho_q W \rho_q^{-1}$, we have $ZS_q = S_q Z$.

2. If W is represented by (5.44), then

$$\begin{aligned} S^q W h &= \pi^q z \pi_- ph = \pi_- q^{-1} \pi_+ q z \pi_- ph = \pi_- p q^{-1} q^{-1} z \pi_- ph \\ &= \pi_- z \pi_- ph = W S^q h. \end{aligned}$$

Conversely, assume $WS^q = S^q W$, and define Z by (5.43). Then we have $ZS_q = S_q Z$. Applying Theorem 5.17, there exists a polynomial $p(z) \in \mathbb{F}[z]$ for which $Z = p(S_q)$. Since $W = \rho_q^{-1} Z \rho_q$, this implies

$$Wh = \rho_q^{-1} Z \rho_q h = q^{-1} q \pi_- q^{-1} pqh = \pi_- ph. \qquad \blacksquare$$

Theorem 5.28. *Given the polynomials $p(z), q(z)$, with $q(z)$ monic. Let $r(z)$ and $s(z)$ be the g.c.d. and the l.c.m. of $p(z)$ and $q(z)$ respectively. Then*

1. *We have the factorizations*

$$\begin{aligned} q(z) &= r(z) q_1(z), \\ p(z) &= r(z) p_1(z), \end{aligned} \tag{5.45}$$

with $p_1(z), q_1(z)$ coprime, as well as the factorizations

$$s(z) = r(z) p_1(z) q_1(z) = p(z) q_1(z) = q(z) p_1(z). \tag{5.46}$$

We have

$$\begin{aligned} \operatorname{Ker} p(S^q) &= X^r, \\ \operatorname{Im} p(S^q) &= X^{q_1}. \end{aligned} \tag{5.47}$$

2. *Let $p(z), q(z) \in \mathbb{F}[z]$, with $q(z)$ nonzero. Then the linear transformation $p(S^q)$ is invertible if and only if $p(z)$ and $q(z)$ are coprime. Moreover, we have*

$$p(S^q)^{-1} = a(S^q), \tag{5.48}$$

where the polynomial $a(z)$ arises out of any solution of the Bezout equation

$$a(z) p(z) + b(z) q(z) = 1. \tag{5.49}$$

$\blacksquare$

5.5 The Chinese Remainder Theorem and Interpolation

The roots of the Chinese remainder theorem are in number theory. However, we interpret it, the underlying ring taken to be $\mathbb{F}[z]$, as an interpolation result.

Theorem 5.29 (Chinese remainder theorem). *Let $q_i(z) \in \mathbb{F}[z]$ be mutually coprime polynomials and let $q(z) = q_1(z) \cdots q_s(z)$. Then, given polynomials $a_i(z)$ such that $\deg a_i < \deg q_i$, there exists a unique polynomial $f(z) \in X_q$, i.e., such that $\deg f < \deg q$, and for which $\pi_{q_i} f = a_i$.*

Proof. The interesting thing about the proof of the Chinese remainder theorem is its use of coprimeness in two distinct ways. One is geometric, the other one is spectral. Let us define $d_j(z) = \prod_{i \neq j} q_i(z)$.

The mutual coprimeness of the $q_i(z)$ implies the direct sum decomposition

$$X_q = d_1 X_{q_1} \oplus \cdots \oplus d_s X_{q_s}. \tag{5.50}$$

This is the geometric use of the coprimeness assumption. The condition $\deg f < \deg q$ is equivalent to $f(z) \in X_q$. Let $f(z) = \sum_{j=1}^{s} d_j(z) f_j(z)$ with $f_j(z) \in X_{q_j}$. Since for $i \neq j$, $q_i(z) \mid d_j(z)$, it follows that in this case $\pi_{q_i} d_j f_j = 0$. Hence

$$\pi_{q_i} f = \pi_{q_i} \sum_{j=1}^{s} d_j f_j = \pi_{q_i} d_i f_i = d_i(S_{q_i}) f_i.$$

For the spectral use of coprimeness, we observe that the pairwise coprimeness of the $q_i(z)$ implies the coprimeness of $d_i(z)$ and $q_i(z)$. In turn, this shows, applying Theorem 5.18, that the module homomorphism $d_i(S_{q_i})$ in X_{q_i} is actually an isomorphism. Hence, there exists a unique $f_i(z)$ in X_{q_i} such that $a_i = d_i(S_{q_i}) f_i$ and $f_i = d_i(S_{q_i})^{-1} a_i$. So $f = \sum_{j=1}^{s} d_j d_i(S_{q_i})^{-1} a_i$ is the required polynomial. Note that the inversion of $d_i(S_{q_i})$ can be done easily, as in Theorem 5.18, using the Euclidean algorithm.

The uniqueness of $f(z)$, under the condition $f(z) \in X_q$, follows from the fact that (5.50) is a direct sum representation. This completes the proof. ∎

5.5.1 Lagrange Interpolation Revisited

Another solution to the Lagrange interpolation problem, introduced in Chapter 2, can be easily derived from the Chinese remainder theorem. Indeed, given distinct numbers $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, n$, the polynomials $q_i(z) = (z - \alpha_i)$ are mutually coprime. We define $q(z) = \Pi_{j=1}^{n} q_j(z)$ and $q_i(z) = \Pi_{j \neq i} q_j(z)$, which leads to the factorizations $q(z) = d_i(z) q_i(z)$. The coprimeness assumption implies the direct sum representation $X_q = d_1 X_{q_1} \oplus \cdots \oplus d_n X_{q_n}$. Thus, any polynomial $f(z) \in X_q$ has a unique representation of the form

$$f(z) = \sum_{j=1}^{n} c_j d_j(z). \tag{5.51}$$

We want to find a polynomial $f(z)$ that satisfies the interpolation conditions $f(\lambda_i) = a_i$, $i = 1, \ldots, n$. Applying the projection π_{q_i} to the expansion (5.51), and noting that $\pi_{q_i} d_j = 0$ for $j \neq i$, we obtain

$$a_i = f(\lambda_i) = \pi_{q_i} f = \pi_{q_i} \sum_{j=1}^{n} c_j d_j(z) = c_i d_i(\lambda_i).$$

Defining the Lagrange interpolation polynomials by $l_i(z) = d_i(z)/d_i(\lambda_i)$, we get $f(z) = \sum_{j=1}^{n} a_j l_j(z)$ for the unique solution in X_q of the Lagrange interpolation problem. Any other solution differs by a polynomial that has a zero at all the points λ_i, hence is divisible by $q(z)$.

5.5.2 *Hermite Interpolation*

We apply now the Chinese remainder theorem to the problem of higher-order interpolation, or **Hermite interpolation**. In Hermite interpolation, which is a generalization of the Lagrange interpolation problem, we prescribe not only the value of the interpolating polynomial at given points, but also the value of a certain number of derivatives, the number of which may differ from point to point.

Specifying the first ν derivatives, counting from zero, of a polynomial $p(z)$ at a point α means that we are given a representation

$$f(z) = \sum_{i=0}^{\nu-1} f_{i,\alpha} (z - \alpha_i)^i + (z - \alpha)^\nu g(z).$$

Of course, since $\deg \sum_{i=0}^{\nu-1} f_{i,\alpha} (z - \alpha_i)^i < \nu$, this means that

$$\deg \sum_{i=0}^{\nu-1} f_{i,\alpha} (z - \alpha)^i = \deg \pi_{(z-\alpha)^\nu} f.$$

Hence we can formulate the **Hermite interpolation problem:** Given distinct $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$, positive integers $\nu_1, \ldots, \nu_k$, and polynomials $f_i(z) = \sum_{j=0}^{\nu_i - 1} f_{j,\alpha_i} (z - \alpha_i)^j$, find a polynomial $f(z)$ such that

$$\pi_{(z-\alpha_i)^{\nu_i}} f = f_i, \qquad i = 1, \ldots, k. \tag{5.52}$$

Proposition 5.30. *There exists a unique solution $f(z)$, of degree $< n = \sum_{i=1}^{k} \nu_i$, to the Hermite interpolation problem. Any other solution of the Hermite interpolation problem is of the form $f(z) + p(z)g(z)$, where $g(z)$ is an arbitrary polynomial and $p(z)$ is given by*

$$p(z) = \Pi_{i=1}^{k} (z - \alpha_i)^{\nu_i}. \tag{5.53}$$

Proof. We apply the Chinese remainder theorem. Obviously, the polynomials $(z-\alpha_i)^{\nu_i}$, $i=1,\ldots,k$ are mutually coprime. Then, with $p(z)$ defined by (5.53), there exists a unique $f(z)$ with $\deg f < n$ for which (5.52) holds.

If $\hat{f}(z)$ is any other solution, then $h(z) = (f(z) - \hat{f}(z))$ satisfies $\pi_{(z-\alpha_i)^{\nu_i}} h = 0$, that is $h(z)$ is divisible by $(z-\alpha_i)^{\nu_i}$. Since these polynomials are mutually coprime, it follows that $p(z) \mid h(z)$, or $h(z) = p(z)g(z)$ for some polynomial $g(z)$. ■

5.5.3 Newton Interpolation

There are situations in which the interpolation data are given to us sequentially. At each stage we solve the corresponding interpolation problem. Newton interpolation is recursive in the sense that the solution at time k is the basis for the solution at time $k+1$.

Definition 5.31. Given $\lambda_i, a_i \in \mathbb{F}$, $i = 0,1,\ldots$, we define the **Newton interpolation problem** NIP(i):
Find polynomials $f_i(z) \in X_{d_i}$, $i \geq 1$, satisfying the interpolation conditions

$$\mathbf{NIP}(i) : f_i(\lambda_j) = a_j, \qquad 0 \leq j \leq i-1. \tag{5.54}$$

Theorem 5.32. *Given $\lambda_i, a_i \in \mathbb{F}$, $i = 0,1,\ldots$, with the λ_i distinct, we define polynomials by*

$$\begin{aligned} q_i(z) &= z - \lambda_i, \\ d_i(z) &= \Pi_{j=0}^{i-1}(z-\lambda_j). \end{aligned} \tag{5.55}$$

Then

1. *We have the factorizations*

$$d_{i+1}(z) = d_i(z)q_i(z), \tag{5.56}$$

   with $d_i(z), q_i(z)$ coprime.
2. *We have the following direct sum decomposition:*

$$X_{d_{i+1}} = X_{d_i} \oplus d_i X_{q_i}. \tag{5.57}$$

3. *Every $f(z) \in X_{d_{i+1}}$ has a unique representation of the form*

$$f(z) = g(z) + cd_i(z), \tag{5.58}$$

   for some $g(z) \in X_{d_i}$ and $c \in \mathbb{F}$.

4. *Let $f_i(z)$ be the solution to the Newton interpolation problem* **NIP**(i)*, then there exists a constant for which*

$$f_{i+1}(z) = f_i(z) + d_i(z)c_i \tag{5.59}$$

is the solution to the Newton interpolation problem **NIP**$(i+1)$ *and it has the representation*

$$f_{i+1}(z) = \sum_{j=0}^{i} c_j d_j(z), \tag{5.60}$$

where

$$c_j = \frac{f_{j+1}(\lambda_j) - f_j(\lambda_j)}{d_j(\lambda_j)}. \tag{5.61}$$

Proof. 1. Follows from (5.55). The coprimeness of $d_i(z)$ and $q_i(z)$ is a consequence of our assumption that the λ_j are distinct.
2. Follows from the factorization (5.56) by applying Proposition 5.6.
3. Follows from the direct sum representation (5.57).
4. In equation (5.59), we substitute the corresponding expression for $f_i(z)$ and proceed inductively to get (5.60).

   Alternatively, we note that $\deg d_i = i$, for $i \geq 0$; hence $\{d_0(z), \ldots, d_i(z)\}$ is a basis for $X_{d_{i+1}}$ and an expansion (5.60) exists.

   Let $f_{i+1}(z)$ be defined by (5.59). Since $d_i(\lambda_j) = 0$ for $j = 0, \ldots, i-1$, it follows that $f_{i+1}(\lambda_j) = a_j$ for $j = 0, \ldots, i-1$. All we need for $f_{i+1}(z)$ to be a solution of **NIP**$(i+1)$ is to choose c_i such that $a_{i+1} = f_{i+1}(\lambda_i) = f_i(\lambda_i) + d_i(\lambda_i)c_i$ or, since $d_i(\lambda_i) \neq 0$, we get (5.61). ■

5.6 Duality

The availability of both polynomial and rational models allows us to proceed with a deeper study of duality. Our aim is to obtain an identification of the dual space to a polynomial model in terms of a polynomial model.

On $\mathbb{F}(z)$, and more generally on $\mathbb{F}((z^{-1}))$, we introduce a bilinear form as follows. Given $f(z) = \sum_{j=-\infty}^{n_f} f_j z^j$ and $g(z) = \sum_{j=-\infty}^{n_g} g_j z^j$, let

$$[f, g] = \sum_{j=-\infty}^{\infty} f_j g_{-j-1}. \tag{5.62}$$

Clearly, the sum in (5.62) is well defined, since only a finite number of summands are nonzero. Given a subspace $M \subset \mathbb{F}(z)$, we let $M^{\perp} = \{f \in \mathbb{F}(z) | [m, f] = 0, \forall m \in M\}$. It is easy to check that $\mathbb{F}[z]^{\perp} = \mathbb{F}[z]$.

We will need the following simple computational rule.

Proposition 5.33. *Let $\phi(z), f(z), g(z)$ be rational functions. Then*

$$[\phi f, g] = [f, \phi g]. \tag{5.63}$$

Proof. With the obvious notation we compute

$$\begin{aligned}[\phi f, g] &= \textstyle\sum_{j=-\infty}^{\infty} (\phi f)_j h_{-j-1} = \sum_{j=-\infty}^{\infty} (\sum_{i=-\infty}^{\infty} \phi_i f_{j-i}) h_{-j-1} \\ &= \textstyle\sum_{i=-\infty}^{\infty} f_{j-i} \sum_{j=-\infty}^{\infty} \phi_i h_{-j-1} = \sum_{k=-\infty}^{\infty} f_k \sum_{j=-\infty}^{\infty} \phi_{j-k} h_{-j-1} \\ &= \textstyle\sum_{k=-\infty}^{\infty} f_k \sum_{i=-\infty}^{\infty} \phi_i h_{-i-k-1} = \sum_{k=-\infty}^{\infty} f_k (\phi h)_{-k-1} = [f, \phi h].\end{aligned}$$

■

Multiplication operators in $\mathbb{F}(z)$ of the form $L_\phi h = \phi h$ are called **Laurent operators**. The function ϕ is called the **symbol** of the Laurent operator.

Before getting the representation of the dual space to a polynomial model, we derive a representation of the dual space to $\mathbb{F}[z]$.

Theorem 5.34. *The dual space of $\mathbb{F}[z]$ can be identified with $z^{-1}\mathbb{F}[[z^{-1}]]$.*

Proof. Clearly, every element $h(z) \in z^{-1}\mathbb{F}[[z^{-1}]]$ defines, by way of the pairing (5.62), a linear functional on $\mathbb{F}[z]$. Conversely, given a linear functional Φ on $\mathbb{F}[z]$, it induces linear functionals ϕ_i on $\mathbb{F}$ by defining, for $\xi \in \mathbb{F}$,

$$\phi_i(\xi) = [z^i \xi, \phi] = \Phi(z^i \xi),$$

and an element $h \in z^{-1}\mathbb{F}[[z^{-1}]]$ is defined by letting $h(z) = \sum_{j=0}^{\infty} \phi_j z^{-j-1}$. It follows that $\Phi(f) = [f, h]$. ■

Point evaluations are clearly linear functionals in $\mathbb{F}[z]$. It is easy to identify the representing functions. In fact, this is an algebraic version of Cauchy's theorem.

Proposition 5.35. *Let $\alpha \in \mathbb{F}$ and $f(z) \in \mathbb{F}[z]$. Then*

$$f(\alpha) = [f, (z-\alpha)^{-1}].$$

Proof. We have $(z-\alpha)^{-1} = \sum_{i=1}^{\infty} \alpha^{i-1} z^{-i}$. So this follows from (5.62), since

$$\left[f, \frac{1}{z-\alpha}\right] = \sum_{i=0}^{n} f_i \alpha^i = f(\alpha).$$

■

Theorem 5.36. *Let $M = d\mathbb{F}[z]$ with $d(z) \in \mathbb{F}[z]$. Then $M^\perp = X^d$.*

Proof. Let $f(z) \in \mathbb{F}[z]$ and $h(z) \in M$. Then

$$0 = [df, h] = [f, dh] = [f, \pi_- dh].$$

But this implies $d(z)h(z) \in X_d$, or $h(z) \in X^d$. ∎

Next, we compute the adjoint of the projection $\pi_d : \mathbb{F}[z] \longrightarrow \mathbb{F}[z]$. Clearly π_d^* is a transformation acting in $z^{-1}\mathbb{F}[[z^{-1}]]$.

Theorem 5.37. *The adjoint of π_d is π^d.*

Proof. Let $f(z) \in \mathbb{F}[z]$ and $h(z) \in z^{-1}\mathbb{F}[[z^{-1}]]$. Then

$$\begin{aligned}[\pi_d f, h] &= [d\pi_- d^{-1} f, h] = [\pi_- d^{-1} f, \tilde{d}h] = [d^{-1} f, \pi_+ \tilde{d}h] \\ &= [f, \tilde{d}^{-1}\pi_+ \tilde{d}h] = [\pi_+ f, \tilde{d}^{-1}\pi_+ \tilde{d}h] = [f, \pi_- \tilde{d}^{-1}\pi_+ \tilde{d}h] \\ &= [f, \pi^{\tilde{d}} h].\end{aligned}$$

∎

Not only are we interested in the study of duality on the level of $\mathbb{F}[z]$ and its dual space $z^{-1}\mathbb{F}[[z^{-1}]]$, but also we would like to study it on the level of the modules X_d and X^d. The key to this study is the fact that if X is a vector space and M a subspace, then $(X/M)^* \simeq M^\perp$.

Theorem 5.38. *Let $d(z) \in \mathbb{F}[z]$ be nonsingular. Then X_d^* is isomorphic to X^d and $S_d^* = S^d$.*

Proof. Since X_d is isomorphic to $\mathbb{F}[z]/d\mathbb{F}[z]$, then X_d^* is isomorphic to $(\mathbb{F}[z]/d\mathbb{F}[z])^*$, which, in turn, is isomorphic to $(d\mathbb{F}[z])^\perp$. However, this last module is X^d. It is clear that under the duality pairing we introduced, we actually have $X_d^* = X^d$. Finally, let $f(z) \in X_d$ and let $h(z) \in X^d$. Then

$$\begin{aligned}[S_d f, h] &= [\pi_d zf, h] = \left[zf, \pi^d h\right], \\ [zf, h] &= [f, zh] = [\pi_+ f, zh], \\ [f, \pi_- zh] &= [f, S_- h] = \left[f, S^d h\right].\end{aligned}$$

Hence, we can identify $X_{\tilde{d}}$ with X_d^* by defining a new pairing

$$\langle f, g \rangle = [d^{-1} f, g] = [f, d^{-1} g] \tag{5.64}$$

for all $f(z), g(z) \in X_d$. ∎

As a direct corollary of Theorem 5.38 we have the following.

Theorem 5.39. *The dual space of X_d under the pairing $\langle\ ,\ \rangle$ introduced in (5.64) is X_d, and moreover, $S_d^* = S_d$.* ∎

With the identification of the polynomial model X_d with its dual space, we can identify some pairs of dual bases.

Proposition 5.40. *1. Let $d(z) = z^n + d_{n-1}z^{n-1} + \cdots + d_0$, and let $\mathscr{B}_{st} = \{1, z, \ldots, z^{n-1}\}$ be the standard and $\mathscr{B}_{co} = \{e_1(z), \ldots, e_n(z)\}$ the control bases respectively of X_d. Then $\mathscr{B}_{co} = \mathscr{B}_{st}^*$, that is, the control and standard bases are dual to each other.*

2. Let $d(z) = \Pi_{i=1}^n (z - \lambda_i)$, with the λ_i distinct. Let $\mathscr{B}_{in} = \{\pi_1(z), \ldots, \pi_n(z)\}$, with $\pi_i(z)$ the Lagrange interpolation polynomials, be the interpolation basis in X_d. Let $\mathscr{B}_{sp} = \{p_1(z), \ldots, p_n(z)\}$ be the spectral basis, with $p_i(z) = \Pi_{j \neq i}(z - \lambda_j)$. Then $\mathscr{B}_{sp}^* = \mathscr{B}_{in}$.

Proof. 1. We use (5.64) to compute

$$\begin{aligned} \langle z^{i-1}, e_j \rangle &= [d^{-1} z^{i-1}, \pi_+ z^{-j} d] = [d^{-1} z^{i-1}, z^{-j} d] \\ &= [z^{i-j-1}, 1] = \delta_{ij}. \end{aligned}$$

2. We use Proposition 5.35 and note that for every $f(z) \in X_d$ we have

$$\langle f, p_i \rangle = [d^{-1} f, p_i] = \left[f, \frac{p_i}{d} \right] = \left[f, \frac{1}{z - \lambda_i} \right] = f(\lambda_i).$$

In particular, $\langle \pi_i, p_j \rangle = \pi_i(\lambda_j) = \delta_{ij}$. ∎

This result explains the connection between the two companion matrices given in Proposition 5.4. Indeed,

$$C_q^\sharp = [S_q]_{st}^{st} = \widetilde{[S_q^*]_{co}^{co}} = \widetilde{[S_q]_{co}^{co}} = \widetilde{C_q^\flat}.$$

Next we compute the change of basis transformations.

Proposition 5.41. *1. Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$,. Then*

$$[I]_{co}^{st} = \begin{pmatrix} q_1 & \cdot & \cdot & q_{n-1} & 1 \\ \cdot & & & \cdot & \\ \cdot & & \cdot & & \\ q_{n-1} & \cdot & & & \\ 1 & & & & \end{pmatrix} \tag{5.65}$$

and

$$[I]_{st}^{co} = \begin{pmatrix} & & & & 1 \\ & & & \cdot & \psi_1 \\ & & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \\ 1 & \psi_1 & \cdot & \cdot & \psi_{n-1} \end{pmatrix}, \tag{5.66}$$

where for $q^\sharp(z) = z^n q(z^{-1})$, $\psi(z) = \psi_0 + \cdots + \psi_{n-1}z^{n-1}$ *is the unique solution, of degree* $< n$, *of the Bezout equation*

$$q^\sharp(z)\psi(z) + z^n\sigma(z) = 1. \tag{5.67}$$

2. *The matrices in (5.65) and (5.66) are inverses of each other.*
3. *Let* $d(z) = \Pi_{j=1}^n(z-\alpha_j)$ *with* $\alpha_1,\ldots,\alpha_n$ *distinct and let* $\mathscr{B}_{sp}, \mathscr{B}_{in}$ *be the corresponding spectral and interpolation bases of* X_d. *Then we have the following change of basis transformations:*

$$[I]_{st}^{in} = \begin{pmatrix} 1 & \alpha_1 & . & . & \alpha_1^{n-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & \alpha_n & . & . & \alpha_n^{n-1} \end{pmatrix}, \tag{5.68}$$

$$[I]_{sp}^{co} = \begin{pmatrix} 1 & . & . & . & 1 \\ \alpha_1 & . & . & . & \alpha_n \\ . & . & . & . & . \\ . & . & . & . & . \\ \alpha_1^{n-1} & . & . & . & \alpha_n^{n-1} \end{pmatrix}, \tag{5.69}$$

$$[I]_{sp}^{in} = \begin{pmatrix} p_1(\alpha_1) & & & \\ & . & & \\ & & . & \\ & & & . \\ & & & p_n(\alpha_n) \end{pmatrix}. \tag{5.70}$$

Proof. 1. Note that if $\psi(z)$ is a solution of (5.67), then necessarily $\psi_0 = 1$. With J the transposition matrix defined in (8.39), we compute

$$J[I]_{st}^{co} = ([I]_{co}^{st}J)^{-1} = (q^\sharp(S_{z^n}))^{-1}.$$

However, if we consider the map S_{z^n}, then

$$\begin{pmatrix} 1 & q_{n-1} & . & . & . & q_1 \\ & . & & . & . & . \\ & & & . & . & . \\ & & & & . & . \\ & & & & & q_{n-1} \\ & & & & & 1 \end{pmatrix} = q^\sharp(S_{z^n}),$$

and its inverse is given by $\psi(S_{z^n})$, where $\psi(z)$ solves the Bezout equation (5.67). This completes the proof.

2. Follows from the fact that $[I]_{co}^{st}[I]_{st}^{co} = I$.
3. The matrix representation for $[I]_{st}^{in}$ has been derived in Corollary 2.36.

The matrix representation for $[I]_{sp}^{co}$ follows from (5.68) by applying duality theory, in particular Theorem 4.38. We can also derive this matrix representation directly, which we proceed to do. For this, we define polynomials $s_1(z), \dots, s_n(z)$ by

$$s_i(z) = e_1(z) + \alpha_i e_2(z) + \cdots + \alpha_i^{n-1} e_n(z).$$

We claim that $s_i(z)$ are eigenfunctions of S_d corresponding to the eigenvalues α_i. Indeed, using equation (5.12) and the fact that $0 = q(\alpha_i) = q_0 + q_1\alpha_i + \cdots + \alpha_i^n$, we have

$$\begin{aligned} S_q s_i &= -q_0 e_n + \alpha_i(e_1 - q_1 e_n) + \cdots + \alpha_i(e_{n-1} - q_{n-1} e_n) \\ &= \alpha_i e_1 + \cdots + \alpha_i^{n-1} e_{n-1} - (q_0 + \cdots + q_{n-1}\alpha_i^{n-1}) e_n \\ &= \alpha_i e_1 + \cdots + \alpha_i^n e_n = \alpha_i s_i(z). \end{aligned}$$

This implies that there exist constants γ_i such that $s_i(z) = \gamma_i p_i(z)$. Since both $s_i(z)$ and $e_i(z)$ are obviously monic, it follows that necessarily $\gamma_i = 1$ and $s_i(z) = p_i(z)$. The equations

$$p_i(z) = e_1(z) + \alpha_i e_2(z) + \cdots + \alpha_i^{n-1} e_n(z)$$

imply the matrix representation (5.69).

Finally, the matrix representation in (5.70) follows from the trivial identities

$$p_i(z) = p_i(\alpha_i)\pi_i(z).$$

■

5.7 Universality of Shifts

We end this chapter by explaining the reason that the polynomial and rational models we introduced can be used so effectively in linear algebra and its applications. The main reason for this is the universality property of shifts. For our purposes, we shall show that any linear transformation in a finite-dimensional vector space $\mathscr{V}$ over the field $\mathbb{F}$ is isomorphic to the compression of the forward shift $S_+ : \mathbb{F}[z]^n \longrightarrow \mathbb{F}[z]^n$, defined in (1.31), to a quotient space, or alternatively, to the restriction of the backward shift $S_- : z^{-1}\mathbb{F}[[z^{-1}]]^n \longrightarrow z^{-1}\mathbb{F}[[z^{-1}]]^n$, defined in (1.32), to an invariant subspace. For greater generality, we will have to stray away from scalar-valued polynomial and rational functions.

Assume we are given a linear transformation A in a finite-dimensional vector space $\mathscr{V}$, over the field $\mathbb{F}$. Without loss of generality, by choosing a matrix representation, we may as well assume that $\mathscr{V} = \mathbb{F}^n$ and that A is a square matrix.

By $\mathbb{F}^n[z]$ we denote the space of vector polynomials with coefficients in $\mathbb{F}^n$, whereas $\mathbb{F}[z]^n$ will denote the space of vectors with polynomial entries. We will use freely the isomorphism between $\mathbb{F}^n[z]$ and $\mathbb{F}[z]^n$ and we will identify the two spaces.

With the linear polynomial matrix $zI - A$ we associate a map $\pi_{zI-A} : \mathbb{F}[z]^n \longrightarrow \mathbb{F}^n$, given by

$$\pi_{zI-A} \sum_{j=0}^{k} \xi_j z^j = \sum_{j=0}^{k} A^j \xi_j. \tag{5.71}$$

The operation defined above can be considered as taking the remainder of a polynomial vector after left division by the polynomial matrix $zI - A$.

Proposition 5.42. *For the map π_{zI-A} defined by (5.71)*

1. *We have π_{zI-A} is surjective.*
2. *We have*
$$\operatorname{Ker} \pi_{zI-A} = (zI - A)\mathbb{F}[z]^n. \tag{5.72}$$
3. *For the map $S_+ : \mathbb{F}[z]^n \longrightarrow \mathbb{F}[z]^n$ defined by*
$$(S_+ f)(z) = zf(z), \tag{5.73}$$
the following diagram is commutative:
$$\begin{array}{ccc} \mathbb{F}[z]^n & \xrightarrow{\pi_{zI-A}} & \mathbb{F}^n \\ \Big\downarrow{\scriptstyle S_+} & & \Big\downarrow{\scriptstyle A} \\ \mathbb{F}[z]^n & \xrightarrow{\pi_{zI-A}} & \mathbb{F}^n \end{array}$$
This implies that
$$Ax = \pi_{zI-A} z \cdot x. \tag{5.74}$$
4. *Given a polynomial $p(z) \in \mathbb{F}[z]$, we have*
$$p(A)x = \pi_{zI-A} p(z)x. \tag{5.75}$$

Proof. 1. For each constant polynomial $x \in \mathbb{F}[z]^n$, we have $\pi_{zI-A}x = x$. The surjectivity follows.

2. Assume $f(z) = (zI - A)g(z)$ with $g(z) = \sum_{i=0}^{k} g_i z^i$. Then

$$f(z) = (zI - A) \sum_{i=0}^{k} g_i z^i = \sum_{i=0}^{k} g_i z^{i+1} - \sum_{i=0}^{k} A g_i z^i.$$

Therefore

$$\pi_{zI-A} f = \sum_{i=0}^{k} A^{i+1} g_i - \sum_{i=0}^{k} A^i A g_i = 0,$$

i.e., $(zI - A)\mathbb{F}[z]^n \subset \operatorname{Ker} \pi_{zI-A}$.

Conversely, assume $f(z) = \sum_{i=0}^{k} f_i z^i \in \operatorname{Ker} \pi_{zI-A}$, that is, $\sum_{i=0}^{k} A^i f_i = 0$. Recalling that $z^i I - A^i = (zI - A)\sum_{j=0}^{i-1} z^{i-1-j} A^j$, we compute

$$\begin{aligned} f(z) &= \textstyle\sum_{i=0}^{k} f_i z^i = \sum_{i=0}^{k} f_i z^i - \sum_{i=0}^{k} A^i f_i \\ &= \textstyle\sum_{i=0}^{k} (z^i I - A^i) f_i = \sum_{i=0}^{k} (zI - A)(\sum_{j=0}^{i-1} z^{i-1-j} A^j) f_i \\ &= (zI - A) \textstyle\sum_{i=0}^{k} (\sum_{j=0}^{i-1} z^{i-1-j} A^j) f_i = (zI - A) g, \end{aligned}$$

so $\operatorname{Ker} \pi_{zI-A} \subset (zI - A)\mathbb{F}[z]^n$; Hence the equality (5.72) follows.

3. To prove the commutativity of the diagram, we compute, with $f(z) = \sum_{i=0}^{k} f_i z^i$,

$$\begin{aligned} \pi_{zI-A} S_+ f &= \pi_{zI-A} z \textstyle\sum_{i=0}^{k} f_i z^i = \pi_{zI-A} \sum_{i=0}^{k} f_i z^{i+1} \\ &= \textstyle\sum_{i=0}^{k} A^{i+1} f_i = A \sum_{i=0}^{k} A^i f_i \\ &= A \pi_{zI-A} f. \end{aligned}$$

4. By linearity, it suffices to prove this for polynomials of the form z^k. We do this by induction. For $k = 1$ this holds by equation (5.74). Assume it holds up to $k - 1$. Using the fact that

$$z \operatorname{Ker}(zI - A) = z(zI - A)\mathbb{F}[z]^n \subset (zI - A)\mathbb{F}[z]^n = \operatorname{Ker}(zI - A),$$

we compute

$$\pi_{zI-A} z^k x = \pi_{zI-A} z \pi_{zI-A} z^{k-1} x = \pi_{zI-A} z A^{k-1} x = A A^{k-1} x = A^k x. \qquad \blacksquare$$

Clearly, the equality $f(z) = (zI - A)g(z) + \pi_{zI-A} f$ can be interpreted as $\pi_{zI-A} f$ being the remainder of $f(z)$ after division by $zI - A$.

As a corollary, we obtain the celebrated Cayley–Hamilton theorem.

Theorem 5.43 (Cayley–Hamilton). *Let A be a linear transformation in an n-dimensional vector space $\mathscr{U}$ over $\mathbb{F}$ and let $d_A(z)$ be its characteristic polynomial. Then*

$$d_A(A) = 0.$$

Proof. By Cramer's rule, we have $d_A(z) I = (zI - A)\operatorname{adj}(zI - A)$; hence we have the inclusion

$$d_A(z)\mathbb{F}[z]^n \subset (zI - A)\mathbb{F}[z]^n.$$

This implies, for each $x \in \mathscr{U}$, that $d_A(A)x = \pi_{zI-A} d(z) x = 0$, so $d_A(A) = 0$. $\blacksquare$

Corollary 5.44. *Let A be a linear transformation in an n-dimensional vector space $\mathscr{U}$ over $\mathbb{F}$. Then its minimal polynomial $m_A(z)$ divides its characteristic polynomial $d_A(z)$.*

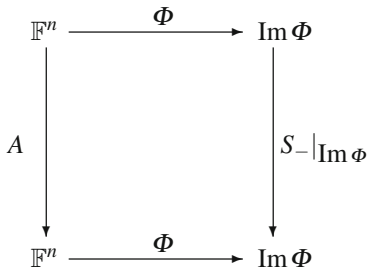
Theorem 5.45. *Let A be a linear transformation in $\mathbb{F}^n$. Then A is isomorphic to S_- restricted to a finite-dimensional S_--invariant subspace of $z^{-1}\mathbb{F}[[z^{-1}]]^n$. Specifically, let $\Phi : \mathbb{F}^n \longrightarrow z^{-1}\mathbb{F}[[z^{-1}]]^n$ be defined by*

$$\Phi\xi = (zI - A)^{-1}\xi, \tag{5.76}$$

and let

$$\mathscr{L} = \operatorname{Im}\Phi. \tag{5.77}$$

Then the following diagram is commutative:



which implies the isomorphism

$$A \simeq S_-|_{\operatorname{Im}\Phi}. \tag{5.78}$$

Proof. Note that the map Φ is injective, hence invertible as a map from $\mathbb{F}^n$ onto $\mathscr{L} = \operatorname{Im}\Phi = \{(zI - A)^{-1}\xi \,|\, \xi \in \mathbb{F}^n\}$. Since

$$(zI - A)^{-1}\xi = \sum_{j=0}^{\infty} A^j \xi z^{-(j+1)}$$

$\mathscr{L}$ is a subspace of $z^{-1}\mathbb{F}[[z^{-1}]]^n$. Since

$$\begin{aligned} S_-\Phi\xi &= S_-(zI - A)^{-1}\xi = \pi_- z(zI - A)^{-1}\xi \\ &= \pi_-(zI - A + A)(zI - A)^{-1}\xi = (zI - A)^{-1}A\xi \\ &= \Phi A\xi, \end{aligned}$$

the S_--invariance of $\mathscr{L}$ follows as well as the isomorphism (5.78). ∎

Remark 5.46. 1. The underlying idea is that the subspace $\operatorname{Im}\Phi$ contains all the information about A. The study of $\mathscr{L} = \operatorname{Im}\Phi$ as an $\mathbb{F}[z]$-module is a tool for the study of the transformation A.

2. The subspace $\mathscr{L} = \operatorname{Im}\Phi$ inherits an $\mathbb{F}[z]$-module structure from $z^{-1}\mathbb{F}[[z^{-1}]]^m$. With $\mathbb{F}^n$ having the $\mathbb{F}[z]$-module structure induced by A, then Φ is clearly an $\mathbb{F}[z]$-module homomorphism.

5.8 Exercises

1. Assume $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$ is a real or complex polynomial. Show that the solutions of the linear, homogeneous differential equation

$$y^{(n)} + q_{n-1}y^{(n-1)} + \cdots + q_0 y = 0$$

   form an n-dimensional space. Use the Laplace transform $\mathscr{L}$ to show that y is a solution if and only if $\mathscr{L}(y) \in X^q$.
2. Let T be a cyclic transformation with minimal polynomial $m_T(z)$ of degree n, and let $p(z) \in \mathbb{F}[z]$. Show that the following statements are equivalent:

   a. The operator $p(T)$ is cyclic.
   b. There exists a polynomial $q(z) \in \mathbb{F}[z]$ such that $(q \circ p)(z) = z \mod (m_T)$.
   c. The map $\Lambda_p : \mathbb{F}[z] \longrightarrow X_{m_T}$ defined by $\Lambda_p(q) = \pi_{m_T}(q \circ p)$ is surjective.
   d. We have $\det(\pi_{kj}) \neq 0$, where $\pi_k = \pi_{m_T}(p^k)$ and $\pi_k(z) = \sum_{j=0}^{n-1} \pi_{kj} z^j$.

3. Assume the minimal polynomial of a cyclic operator T factors into linear factors, i.e., $m_T(z) = \prod (z-\lambda_i)^{\nu_i}$ with the λ_i distinct. Show that $p(T)$ is cyclic if and only if $\lambda_i \neq \lambda_j$ implies $p(\lambda_i) \neq p(\lambda_j)$ and $p'(\lambda_i) \neq 0$ whenever $\nu_i > 1$.
4. Show that if $\pi(z) \in \mathbb{F}[z]$ is irreducible and $\deg p$ is a prime number, then $p(S_\pi)$ is either scalar or a cyclic transformation.
5. Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$ and let $C_q^\sharp$ be its companion matrix. Let $f(z) \in X_q$ with $f(z) = f_0 + \cdots + f_{n-1}z^{n-1}$. We put

$$\mathbf{f} = \begin{pmatrix} f_0 \\ . \\ . \\ . \\ f_{n-1} \end{pmatrix} \in \mathbb{F}^n.$$

   Show that $f(C_q) = \left(\mathbf{f}, C_q\mathbf{f}, \ldots, C_q^{n-1}\mathbf{f}\right)$.
6. Given a linear transformation A in $\mathscr{V}$, a vector $x \in \mathscr{V}$, and a polynomial $p(z)$, define the **Jacobson chain matrix** by

$$C_m(p,x,A) = (x, Ax, \ldots, A^{r-1}x, \ldots, p(A)^{m-1}x, p(A)^{m-1}Ax, \ldots, p(A)^{m-1}A^{r-1}x).$$

   Prove the following

   a. Let $E = C_m(p, 1, C_{p^m})$ then E is a solution of $C_{p^m}E = EH(p^m)$.
   b. Every solution X of $C_{p^m}X = XH(p^m)$ is of the form $X = f(C_{p^m})E$ for some polynomial $f(z)$ of degree $< m \deg p$.

7. Let $p(z) = z^m + p_{m-1}z^{m-1} + \cdots + p_0$ and $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$. Let $H(p,q) = \begin{pmatrix} C_p^\sharp & 0 \\ N & C_q^\sharp \end{pmatrix}$, where N is the $n \times m$ matrix whose only nonzero element is $N_{1m} = 1$.
Show that the general solution to the equation $C_{pq}^\sharp X = XH(pq)$ is of the form $X = f(C_{pq}^\sharp)K$, where $\deg f < \deg p + \deg q$ and

$$K = \begin{pmatrix} 1 & & p_0 & & \\ & \ddots & & \ddots & \\ & & 1 & p_{m-1} & \ddots \\ & & & 1 & p_0 \\ & & & & \ddots \\ & & & & 1 \end{pmatrix}.$$

8. Let C be the circulant matrix

$$C = \operatorname{circ}(c_0, \ldots, c_{n-1}) = \begin{pmatrix} c_0 & c_{n-1} & \ldots & c_1 \\ c_1 & c_0 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ldots & c_{n-1} \\ c_{n-1} & \ldots & c_1 & c_0 \end{pmatrix}.$$

Show that $\det(C) = \prod_{i=1}^n (c_0 + c_1\zeta_i + \cdots + c_{n-1}\zeta_i^{n-1})$, where $1 = \zeta_1, \ldots, \zeta_n$ are the distinct nth roots of unity, that is, the zeros of $z^n - 1$.
9. Prove the following **barycentric representation** for the minimal-degree polynomial satisfying the interpolation constraints $f(\lambda_i) = a_i$, $i = 0, \ldots, n-1$, namely

$$f(z) = \begin{cases} \dfrac{\sum_{i=0}^{n-1} \dfrac{w_j}{z - \lambda_j} a_i}{\sum_{i=0}^{n-1} \dfrac{w_j}{z - \lambda_j}}, & z \neq \lambda_j, j = 0, \ldots, n-1, \\ a_i, & z = \lambda_j, j = 0, \ldots, n-1. \end{cases}$$

Here $w_i = \frac{1}{\Pi_{j \neq i}(\lambda_i - \lambda_j)}$.

5.9 Notes and Remarks

Shift operators are a cornerstone of modern operator theory. As we explained in Section 5.7, their interest lies in their universality properties, an extremely important fact, first pointed out in Rota (1960). This observation is the key insight to the use of shift operators in modeling linear transformations as well as, more generally, linear systems. The fact that any linear transformation T acting in a finite-dimensional vector space is isomorphic to either a compression of S_+ or to a restriction of S_- has far-reaching implications. This leads to the use of functional models, in our case polynomial and rational models, in the study of linear transformations. The reason for the great effectiveness of the use of functional models can be traced back to the compactness of the polynomial notation as well as that the polynomial setting provides a richer algebraic language, with terms as zeros, factorizations, ideals, homomorphisms, and modules conveniently at hand. Of the two classes of models we employ, the polynomial models emphasize the arithmetic properties of the transformation, whereas the module structure of rational models is completely determined by the geometry of the model space. In other words, in the case of polynomial models, the space is simple but the tranformation is complicated. On the other hand, for rational models, the space is complicated but the transformation, i.e., the restricted backward shift, is simple.

The material in this chapter is mostly standard. The definitions given in Equations (5.71) and (5.75) have far reaching implications. They can be generalized, replacing $zI - A$ by an arbitrary nonsingular polynomial matrix, thus leading to the theory of polynomial models, initiated in Fuhrmann (1976). This is the algebraic counterpart of the functional models used so effectively in operator theory, e.g., Sz.-Nagy and Foias (1970).

The results and methods presented in this chapter originate in Fuhrmann (1976) and a long series of follow-up articles. In an analytic, infinite-dimensional setting, Nikolskii (1985) is a very detailed study of shift operators. The central results of this chapter are two theorems. Theorem 5.17 characterizes the maps intertwining the shifts of the form S_q, while Theorem 5.18 studies their invertibility properties in terms of factorizations and polynomial coprimeness. These results generalize to the use of shifts of higher multiplicity, a generalization that requires the use of the algebra of polynomial matrices, which is beyond the scope of the present book.

The analytic analogues of these result are the commutant lifting theorem, see Sarason (1967) and Sz.-Nagy and Foias (1970), and a corresponding spectral mapping theorem; see Fuhrmann (1968a,b). We shall return to these topics in Chapter 11.

The kernel representation (5.37) of rational models has far-reaching generalizations. With the scalar polynomial replaced by a rectangular polynomial matrix, it is the basic characterizations of behaviors in the behavioral approach to linear systems initiated in Willems (1986, 1989, 1991). In this connection, see also Fuhrmann (2002).

Theorem 5.17 can be extended to the characterization of intertwining maps between two polynomial, or rational, models. We will return to this in Chapter 8, where we give a different proof using tensor products and Bezoutians.

Theorem 5.18 highlights the importance of coprimenes and the Bezout equation. Since the Bezout equation can be solved using the Euclidean algorithm, this brings up the possibility of a recursive approach to inversion algorithms for structured matrices. This will be further explored in Chapter 8. The Bezout equation reappears, this time over the ring $\mathbf{RH}_+^\infty$, in Chapters 11 and 12.

Companion matrices of the polynomial $q(z)$ appear in Krull's thesis. The particularly musical notation for the matrices $C_q^\sharp, C_q^\flat$, defined in Proposition 5.4, was introduced by Kalman.

Section 5.3 on circulant matrices is based on results taken from Davis (1979).

The Chinese remainder theorem has its roots in the problem of solving simultaneous congruences arising in modular arithmetic. It appears in a third-century mathematical treatise by Sun Tzu. For a very interesting account of the history of early non-European mathematics, see Joseph (2000). In working over the ring of polynomials $\mathbb{F}[z]$ rather than over the integes $\mathbb{Z}$, it can be applied to interpolation problems. The possibility of applying the Chinese remainder theorem to interpolation problems is not mentioned in many of the classic algebra texts such as van der Waerden, Lang, Mac Lane and Birkhoff or even Gantmacher (1959). Of these monographs, van der Waerden (1931) stands out inasmuch as it discusses both the Lagrange and the Newton interpolation formulas. In connection to interpolation, the following quotation, from Schoenberg (1987), may be of interest: "Sometime in the 1950's the late Hungarian-Swedish mathematican Marcel Riesz visited the University of Pennsylvania and told us informally that the Chinese remainder theorem (1) can be thought of as an analogue of the interpolation by polynomials."

Cayley, in a letter to Sylvester, actually stated a more general version of the Cayley–Hamilton theorem. That version says that if the square matrices A and B commute and $f(x,y) = \det(xA - yB)$, then $f(B,A) = 0$. This was generalized in Livsic (1983).

http://www.springer.com/978-1-4614-0337-1

A Polynomial Approach to Linear Algebra

Fuhrmann, P.A.

2012, XVI, 411 p., Softcover
ISBN: 978-1-4614-0337-1