CYCLIC MATRICES AND POLYNOMIAL INTERPOLATION OVER DIVISION RINGS

VLADIMIR BOLOTNIKOV

Abstract. As is well known, any complex cyclic matrix $A$ is similar to the unique companion matrix associated with the minimal polynomial of $A$. On the other hand, a cyclic matrix over a division ring $F$ is similar to a companion matrix of a polynomial which is defined up to polynomial similarity. In this paper we study more rigid canonical forms by embedding a given cyclic matrix over a division ring $F$ into a controllable or an observable pair. Using the characterization of ideals in $F[z]$ in terms of controllable and observable pairs we consider ideal interpolation schemes in $F[z]$ which merge into a polynomial interpolation problems containing both left and right interpolation conditions.

1. Introduction

Given a complex matrix $A \in \mathbb{C}^{n \times n}$ and a vector $v \in \mathbb{C}^n$, the sets

$$\mathbb{I}_A := \{ p \in \mathbb{C}[z] : p(A) = 0 \} \quad \text{and} \quad \mathbb{I}_{A,v} := \{ p \in \mathbb{C}[z] : p(A)v = 0 \}$$

are ideals in the ring $\mathbb{C}[z]$ of complex polynomials; their respective (monic) generators $\mu_A$ and $\mathbb{P}_{A,v}$ are called the minimal polynomial of the matrix $A$ and the minimal polynomial of the pair $(A,v)$, respectively. As $\mathbb{I}_A \subseteq \mathbb{I}_{A,v}$, it follows that $\mathbb{P}_{A,v}$ divides $\mu_A$.

A matrix $A \in \mathbb{C}^{n \times n}$ is called cyclic if there exists a (cyclic) vector $v \in \mathbb{C}^n$ such that $\text{span}\{ v, Av, \ldots, A^{n-1}v \} = \mathbb{C}^n$, i.e., the controllability matrix

$$\mathbb{C}_{A,v} = \begin{bmatrix} v & Av & \ldots & A^{n-1}v \end{bmatrix}$$

is invertible (equivalently, $\deg(\mathbb{P}_{A,v}) = n$). In this case, we say that the pair $(A,v)$ is controllable. Cyclic matrices and controllable pairs can be characterized in interpolation terms as follows.

Proposition 1.1. (1) The matrix $A \in \mathbb{C}^{n \times n}$ is cyclic if and only if for any $B \in \mathbb{C}^{n \times n}$ commuting with $A$, there is an $f \in \mathbb{C}[z]$ such that $f(A) = B$.

(2) The pair $(A,v)$ with $A \in \mathbb{C}^{n \times n}$ is controllable if and only if for any $b \in \mathbb{C}^n$, there is an $f \in \mathbb{C}[z]$ such that $f(A)v = b$.

The first objective of this paper is to study possible extensions of these results as well as of some other characterizations of complex cyclic matrices recalled in Proposition 1.2 below to the non-commutative setting of a division ring $F$. To fix notation, we let $e_j$ to denote the $j$-th column of the
$n \times n$ identity matrix $I_n = [e_1 \ldots e_n]$ (occasionally writing $e_{j,n}$ if the dimension is not clear from the context). We will use notation
\[ A_n(z) := [1 \, z \, \cdots \, e_n] \] (1.1)
to identify a polynomial $g(z) = A_n(z)g$ with the column $g$ of its coefficients. In terms of this notation, we recall the companion matrix $C(f)$ of a monic polynomial $f \in \mathbb{C}[z]$:
\[ C(f) = [e_2 \, e_3 \, \ldots \, e_n \, -f] \] if $f(z) = z^n + A_n(z)f$. (1.2)

**Proposition 1.2.** Given a matrix $A \in \mathbb{C}^{n \times n}$, the following are equivalent:

1. $A$ is cyclic.
2. $\deg \mu_A = n$, where $\mu_A$ is the minimal polynomial of $A$.
3. $A$ is similar to a (unique) companion matrix (which is $C(\mu_A)$).
4. $A$ is similar to a two-diagonal matrix
\[ \Gamma = \begin{bmatrix}
\gamma_1 & 0 & \cdots & 0 \\
1 & \gamma_2 & \ddots & \vdots \\
& \ddots & \ddots & 0 \\
0 & & 1 & \gamma_n
\end{bmatrix} \] (1.3)
with diagonal entries equal to zeros (roots) of the polynomial $\mu_A$.
5. The Jordan form of $A$ contains only one Jordan block corresponding to each eigenvalue (which is a zero of $\mu_A$).

If this is the case, then $P_{A,v}(z) = \mu_A(z) = \det(zI_n - A)$ for any cyclic vector $v$ of $A$.

In the division-ring setting, it is still true that the matrices (1.2) and (1.3) are cyclic. The rest requires certain adjustments. First, the minimal polynomial $\mu_A$ (more precisely, left and right minimal polynomials; see (2.4)) is not similarity invariant; besides, simple examples show that its degree can be different from the dimension of $A$. The polynomial $P_{A,v}$ seems to be more suitable as it is invariant under similarity of controllable pairs (see Definition 2.4 below) and the equality $\deg(P_{A,v}) = n$ is equivalent to $v$ be a (right) cyclic vector for $A$. Although different cyclic vectors $v, v'$ of $A$ may lead to different minimal polynomials $P_{A,v}$ and $P_{A,v'}$, the latter polynomials are similar: $P_{A,v} \approx P_{A,v'}$ (see Section 4.1 for the precise definition). Thus, in a context that does not distinguish similar polynomials, we may write $P_A$ rather than $P_{A,v}$. Then the well-known fact that two cyclic complex matrices are similar if and only if their minimal polynomials are equal, extends to the non-commutative setting as follows: **cyclic matrices** $A, A' \in \mathbb{F}^{n \times n}$ are similar if and only if $P_A \approx P_{A'}$.

However, there is a more rigid extension in terms of controllable pairs: **two controllable pairs** $(A, v)$ and $(A', v')$ are similar if and only if $P_{A,v} = P_{A',v'}$ (see Theorem 3.3 below).

In a similar manner, part (3) in Proposition 1.2 extends to $\mathbb{F}$-setting in two ways:
(1) A cyclic matrix $A \in \mathbb{F}^{n \times n}$ is similar to the companion matrix $C(P_A)$.

(2) For a fixed cyclic vector $v$ of $A$, the controllable pair $(A, v)$ is similar to a unique pair of the form $(C(f), e_1)$ (with $f = P_A(v)$).

A related result (Theorem 4.4) describes the similarity class of a given polynomial $p \in \mathbb{F}[z]$ in terms of cyclic vectors of the companion matrix $C(p)$. Similarity of a cyclic matrix to two-diagonal or block-diagonal matrices is discussed in Section 4.2.

Our second goal is to study interpolation problems of Hermite-Lagrange type in $\mathbb{F}[z]$. In contrast to the commutative case, polynomials over a division ring can be evaluated on the left and on the right. Simple examples are given by left and right Lagrange interpolation problems where one is seeking an $f \in \mathbb{F}[z]$ with prescribed left (right) values at given points. Solution sets for homogeneous problems are right (left) ideals in $\mathbb{F}[z]$, and under mild the assumption that the set of interpolation nodes is $P$-independent (see Definition 3.8) the problems have unique low-degree solutions. In Section 5 we consider more general interpolation problems with interpolation conditions given in terms of left and right tangential evaluation calculi induced by respectively, controllable and observable pairs. In Section 6 we consider the combined (two-sided) problem containing both left and right interpolation conditions. The solution set of the homogeneous version of this problem is the intersection of a left and a right ideal in $\mathbb{F}[z]$, while the nonhomogeneous problem may have no solutions as well as multiple low-degree solutions. Two-sided Lagrange problem is presented in Section 6 as an illustrative particular case.

2. Preliminaries

Given a division ring $\mathbb{F}$, let $\mathbb{F}[z]$ denote the ring of polynomials in one formal variable $z$ which commutes with coefficients from $\mathbb{F}$. Since the division algorithm holds in $\mathbb{F}[z]$ on either side, any ideal (left or right) in $\mathbb{F}[z]$ is principal. We will write $(f)_r$ and $(f)_\ell$ for the right and the left ideal generated by $f \in \mathbb{F}[z]$ dropping the subscript if the ideal is two-sided (i.e., left and right simultaneously). The intersection of two left (right) ideals is a left (right) ideal; the least right and left common multiples $lrcm(f, g)$ and $llcm(f, g)$ of two monic polynomials $f, g \in \mathbb{F}[z]$ are defined as generators of the respective ideals

$$
(f)_r \cap (g)_r = (lrcm(f, g))_r \quad \text{and} \quad (f)_\ell \cap (g)_\ell = (llcm(f, g))_\ell. \quad (2.1)
$$

If we let $Z_{\mathbb{F}}$ denote the center of $\mathbb{F}$, then $Z_{\mathbb{F}[z]} = Z_{\mathbb{F}}[z]$, and consequently, any ideal generated by an element of $Z_{\mathbb{F}}[z]$ is two-sided. The converse is also true: the generator of any two-sided ideal of $\mathbb{F}[z]$ is in $Z_{\mathbb{F}}[z]$. Indeed, the left and the right (monic) generators of the ideal must be multiples of each other and therefore, they coincide. On the other hand if $p$ is a left and right generator, then it commutes with each $\alpha \in \mathbb{F}$ which implies that its coefficients are in $Z_{\mathbb{F}}$; see e.g., [5, Proposition 2.2.2] for details.
2.1. Minimal polynomials. Any polynomial \( f \in \mathbb{F}[z] \) can be evaluated at a matrix \( A \in \mathbb{F}^{n \times n} \) on the left and on the right (by interpreting \( \mathbb{F}^{n \times n} \) as an \( \mathbb{F} \)-bimodule) as follows:

\[
f^{\text{le}}(A) = \sum A^j f_j \quad \text{and} \quad f^{\text{re}}(A) = \sum f_j A^j, \quad \text{if} \quad f(z) = \sum f_j z^j. \tag{2.2}
\]

Since for any \( f, g \in \mathbb{F}[z] \) and \( A \in \mathbb{F}^{n \times n}, \)

\[
(gf)^{\text{le}}(A) = \sum A^j g^{\text{le}}(A)f_j \quad \text{and} \quad (fg)^{\text{re}}(A) = \sum f_j g^{\text{re}}(A)A^j, \tag{2.3}
\]

it follows that \((gf)^{\text{le}}(A) = 0\) and \((fg)^{\text{re}}(A) = 0\) whenever \(g^{\text{le}}(A) = 0\) and \(g^{\text{re}}(A) = 0\), respectively. Hence, the sets

\[
\mathbb{I}_{A,r} := \{ p \in \mathbb{F}[z] : p^{\text{le}}(A) = 0 \} = \langle \mu_{A,r} \rangle_r,
\]

\[
\mathbb{I}_{A,l} := \{ p \in \mathbb{F}[z] : p^{\text{re}}(A) = 0 \} = \langle \mu_{A,l} \rangle_l \tag{2.4}
\]

are respectively, a right and a left ideal in \( \mathbb{F}[z] \); their generators \( \mu_{A,r} \) and \( \mu_{A,l} \) will be called the **left** and the **right minimal polynomial** of \( A \), respectively.

If \( \mathbb{I}_{A,r} \) contains a non-zero polynomial \( p \in Z_\mathbb{F}[z] \) such that \( p(A) = 0 \), then the set of all such polynomials form the maximal two-sided ideal contained in \( \mathbb{I}_{A,r} \cap \mathbb{I}_{A,l} \) and its generator \( \mu_A \in Z_\mathbb{F}[z] \) is called the **minimal (central) polynomial** of \( A \).

**Remark 2.1.** If \( A = \alpha \in \mathbb{F} \), then evaluations (2.2) amount to left and right “point” evaluations of \( f \) at \( \alpha \):

\[
f^{\text{le}}(\alpha) = \sum \alpha^j f_j \quad \text{and} \quad f^{\text{re}}(\alpha) = \sum f_j \alpha^j. \tag{2.5}
\]

An element \( \alpha \in \mathbb{F} \) is called a **left** or **right zero** of a polynomial \( f \in \mathbb{F}[z] \) if \( f^{\text{le}}(\alpha) = 0 \) or \( f^{\text{re}}(\alpha) = 0 \), respectively. The ideals \( \mathbb{I}_{A,r} \) and \( \mathbb{I}_{A,l} \) (2.4) of polynomials vanishing at \( \alpha \) on the left and on the right respectively, are generated by the linear monic polynomial \( \mu_{A,l} = \mu_{A,r} = \rho_{\alpha} \), where

\[
\rho_{\alpha}(z) := z - \alpha \quad (\alpha \in \mathbb{F}). \tag{2.6}
\]

In other words,

\[
f^{\text{le}}(\alpha) = 0 \iff f \in \langle \rho_{\alpha} \rangle_r \quad \text{and} \quad f^{\text{re}}(\alpha) = 0 \iff f \in \langle \rho_{\alpha} \rangle_l.
\]

The existence of the minimal central polynomial \( \mu_\alpha \in Z_\mathbb{F}[z] \) is the definition of \( \alpha \) being algebraic over \( Z_\mathbb{F} \).

Given a polynomial \( f \in \mathbb{F}[z] \) and given a matrix \( A \in \mathbb{F}^{n \times n} \) and vectors \( v \in \mathbb{F}^{n \times 1} \) and \( u \in \mathbb{F}^{1 \times n} \), one can apply evaluations (2.2) to polynomials \( vf \in \mathbb{F}[z]^{n \times 1} \) and \( fu \in \mathbb{F}[z]^{1 \times n} \) as follows:

\[
(vf)^{\text{le}}(A) = \sum A^j vf_j \quad \text{and} \quad (fu)^{\text{re}}(A) = \sum f_j u A^j. \tag{2.7}
\]

Due to equalities

\[
(vgf)^{\text{le}}(A) = \sum A^j \cdot (vf)^{\text{le}}(A) \cdot f_j = ((vg)^{\text{le}}(A)f)^{\text{le}}(A), \tag{2.8}
\]

\[
(fgu)^{\text{re}}(A) = \sum f_j \cdot (gu)^{\text{re}}(A) \cdot A^j = (fgu)^{\text{re}}(A)^{\text{re}}(A),
\]
holding for all \( f, g \in \mathbb{F}[z] \), the sets
\[
\mathbb{I}_{A,v} := \{ p \in \mathbb{F}[z] : (vp)^{\ell}(A) = 0 \} = \langle \mathfrak{P}_{A,v} \rangle_{\ell},
\]
\[
\mathbb{I}_{u,A} := \{ p \in \mathbb{F}[z] : (pu)^{r}(A) = 0 \} = \langle \mathfrak{P}_{u,A} \rangle_{r}
\]
are respectively a right and a left ideal in \( \mathbb{F}[z] \); their generators are called the minimal polynomials of the input pair \((A, v)\) and of the output pair \((u, A)\), respectively.

Straightforward calculations show that for any \( A \in \mathbb{F}^{n \times n}, \ v \in \mathbb{F}^{n \times 1}, \ u \in \mathbb{F}^{1 \times n}, \) and \( f \in \mathbb{F}[z], \)
\[
v f(z) = (z I_n - A) \cdot (L_A(vf))(z) + (vf)^{\ell}(A),
\]
\[
f(z) u = (R_A(fu))(z) \cdot (z I_n - A) + (fu)^{r}(A),
\]
where \((vf)^{\ell}(A)\) and \((fu)^{r}(A)\) are defined as in (2.7) and where \( L_A(vf) \) and \( R_A(fu) \) are vector polynomials given by
\[
L_A(vf) = \sum_{j+k=0}^{\deg f-1} A^j v f_{k+j+1} z^k, \quad R_A(fu) = \sum_{j+k=0}^{\deg f-1} f_{k+j+1} u A^j z^k. \tag{2.11}
\]

**Remark 2.2.** Representations (2.10) are unique in the following sense: if
\[
v f(z) = (z I_n - A)G(z) + b \quad \text{for some} \quad G \in \mathbb{F}^{n \times 1}[z], \quad b \in \mathbb{F}^{1 \times n},
\]
then necessarily (by comparing the coefficients in the polynomial identity above), \( G = L_A(vf) \) and \( b = (vf)^{\ell}(A) \). The second representation in (2.10) is unique in a similar sense.

Ideals (2.9) can be characterized in terms of evaluations (2.8) as follows.

**Proposition 2.3.** Given \( A \in \mathbb{F}^{n \times n}, \ v \in \mathbb{F}^{n \times 1}, \ u \in \mathbb{F}^{1 \times n} \) and \( f \in \mathbb{F}[z], \)
\( (vf)^{\ell}(A) = 0 \) if and only if \( vf = (z I_n - A) \cdot G \) for some \( G \in \mathbb{F}^{n \times 1}[z] \).
\( (fu)^{r}(A) = 0 \) if and only if \( fu = H \cdot (z I_n - A) \) for some \( H \in \mathbb{F}^{1 \times n}[z] \).

**Proof.** The proof is the same as in the scalar-valued case [7, Theorem 1]. The “only if” parts follow from equalities (2.10). Conversely, comparing the coefficients in the polynomial identity
\[
v f(z) = \sum_{i=0}^{k} v f_k z^k = (z I_n - A) \cdot \sum_{i=0}^{k-1} G_i z^i = (z I_n - A)G(z)
\]
leads us to equalities
\[
v f_0 = -AG_0, \quad v f_k = G_{k-1} \quad \text{and} \quad v f_j = G_{j-1} - AG_j
\]
for \( j = 1, \ldots, k - 1 \), which in turn imply
\[
(vf)^{\ell}(A) = \sum_{j=0}^{k} A^j v f_j = -AG_0 + \sum_{j=1}^{k-1} A^j(G_{j-1} - AG_j) + A^k G_{k-1} = 0.
\]
The proof of part (2) is similar. □
Definition 2.4. Let us say that two input pairs \((A, \mathbf{v})\) and \((A', \mathbf{v}')\) (two output pairs \((\mathbf{u}, A)\) and \((\mathbf{u}', A')\)) are similar if \(A' = TAT^{-1}\) and \(\mathbf{v}' = T\mathbf{v}\) for some invertible matrix \(T \in \mathbb{F}^{n \times n}\).

The next observation follows from Definition 2.4 and formulas (2.7), (2.9).

Remark 2.5. (1) If the input pairs \((A, \mathbf{v})\) and \((A', \mathbf{v}')\) are similar with the similarity matrix \(T\), then \((\mathbf{v}'f)^{\mathbf{e}}(A') = T \cdot (\mathbf{v}f)^{\mathbf{e}}(A)\) and hence \(\mathbb{I}_{A, \mathbf{v}} = \mathbb{I}_{A', \mathbf{v}'}\) and \(\mathcal{P}_{A, \mathbf{v}} = \mathcal{P}_{A', \mathbf{v}'}\).

(2) If the output pairs \((\mathbf{u}, A)\) and \((\mathbf{u}', A')\) are similar, then \((f\mathbf{u})^{\mathbf{e}}(A) = (f\mathbf{u}')^{\mathbf{e}}(A') \cdot T\) and hence \(\mathbb{I}_{\mathbf{u}, A} = \mathbb{I}_{\mathbf{u}', A'}\) and \(\mathcal{P}_{\mathbf{u}, A} = \mathcal{P}_{\mathbf{u}', A'}\).

2.2. Explicit formulas for minimal polynomials. To compute \(\mathcal{P}_{\mathbf{u}, A}\), we first find the least integer \(d\) such that the vectors \(\mathbf{v}, A\mathbf{v}, \ldots, A^d\mathbf{v}\) are (right) linearly dependent and then conclude from the relation

\[
A^d\mathbf{v} + A^{d-1}\mathbf{v}b_{d-1} + \ldots + A\mathbf{v}b_1 + \mathbf{v}b_0 = 0
\]

that

\[
\mathcal{P}_{\mathbf{u}, A}(z) = z^d + \sum_{k=0}^{d-1} z^k b_k.
\]

The construction of \(\mathcal{P}_{\mathbf{u}, A}\) is similar. As for the minimal polynomials \(\mu_{A, \ell}\) and \(\mu_{A, r}\) in (2.4), let us observe that

\[
\mu_{A, \ell} = \text{lrcm}\{\mathcal{P}_{A, e_k} : 1 \leq k \leq n\}, \quad \mu_{A, r} = \text{l lcm}\{\mathcal{P}_{e_k, A}^\top : 1 \leq k \leq n\},
\]

(2.13)

where \(e_1, \ldots, e_n\) denote the columns in the identity matrix \(I_n\). Indeed, since both 0 and 1 belong to the center of \(\mathbb{F}\), it follows from (2.2) and (2.7) that

\[
(e_k f)^{\mathbf{e}}(A) = f^{\mathbf{e}}(A)e_k \quad \text{for} \quad k = 1, \ldots, n,
\]

so that \(\mathbb{I}_{A, r} = \bigcap_{k=1}^n \mathbb{I}_{A, e_k}\), by (2.4) and (2.9). Writing the latter equality in terms of generators we get the first equality in (2.13); the second equality follows similarly. We illustrate the above recipe by computing minimal polynomials of the two-diagonal matrix \(\Gamma\) as in (1.3).

Proposition 2.6. Let \(\Gamma = \Gamma_\gamma \in \mathbb{F}^{n \times n}\) be of the form

\[
\Gamma = \Gamma_\gamma = [\delta_{i,j}\gamma_j + \delta_{i-1,j}]_{i,j=1}^n, \quad \gamma = (\gamma_1, \ldots, \gamma_n) \subset \mathbb{F}^n,
\]

(2.14)

(\(\delta_{i,j}\) is the Kronecker symbol) and let \(e_k\) be the \(k\)-th column of \(I_n\). Then

\[
\mathcal{P}_{\Gamma, e_k} = \rho_{\gamma_k} \rho_{\gamma_{k+1}} \cdots \rho_{\gamma_n} \quad \text{and} \quad \mathcal{P}_{e_k, \Gamma} = \rho_{\gamma_1} \rho_{\gamma_2} \cdots \rho_{\gamma_k}
\]

(2.15)

for all \(k = 1, \ldots, n\), and consequently, the left and right minimal polynomials of \(\Gamma\) are given by

\[
\mu_{\Gamma, \ell} = \text{lrcm}\{\rho_{\gamma_1} \cdots \rho_{\gamma_n}, \rho_{\gamma_2} \cdots \rho_{\gamma_n} \cdots, \rho_{\gamma_{n-1}} \rho_{\gamma_n}, \rho_{\gamma_1}\},
\]

\[
\mu_{\Gamma, r} = \text{l lcm}\{\rho_{\gamma_1} \cdots \rho_{\gamma_n}, \rho_{\gamma_1} \cdots \rho_{\gamma_{n-1}}, \rho_{\gamma_1} \rho_{\gamma_2}, \rho_{\gamma_1}\}.
\]

(2.16)
Proof. We first observe from (2.14) that
\[ \Gamma e_j = e_j \gamma_j + e_{j+1} \quad (j = 1, \ldots, n - 1) \quad \text{and} \quad \Gamma e_n = e_n. \tag{2.17} \]
Therefore, for any fixed \( k \leq n \), the vectors \( e_k, \Gamma e_k, \ldots, \Gamma^{n-k} e_k \) are right linearly independent. By the recipe (2.12), it suffices to find a monic polynomial \( f \) with \( \deg f = n - k \) subject to condition \((e_k f)^\ell(\Gamma) = 0\) to claim that \( P e_k = f. \) We next show that \( f = \rho_{\gamma_k} \rho_{\gamma_{k+1}} \cdots \rho_{\gamma_n} \) is such a polynomial. To this end, we write equalities (2.17) in terms of the left evaluation (2.8) as
\[ (e_j \rho_{\gamma_j})^\ell(\Gamma) = e_{j+1} \quad (j = 1, \ldots, n - 1) \quad \text{and} \quad (e_n \rho_{\gamma_n})^\ell(\Gamma) = 0. \]
Upon making use of the first formula in (2.8) and taking into account the latter equalities for \( j = k, \ldots, n \) we get
\[ (e_k \rho_{\gamma_k} \rho_{\gamma_{k+1}} \cdots \rho_{\gamma_n})^\ell(\Gamma) = ((e_k \rho_{\gamma_k})^\ell(\Gamma) \rho_{\gamma_{k+1}} \cdots \rho_{\gamma_n})^\ell(\Gamma) \]
\[ = (e_{k+1} \rho_{\gamma_{k+1}} \cdots \rho_{\gamma_n})^\ell(\Gamma) \]
\[ = \cdots = (e_n \rho_{\gamma_n})^\ell(\Gamma) = 0, \]
which verifies the first part in (2.15). The second part is verified similarly via writing relations
\[ e_j^\top \Gamma = e_j \gamma_j + e_{j-1} \quad (j = 2, \ldots, n) \quad \text{and} \quad e_1 \Gamma = \gamma_1 e_1 \]
in terms of the evaluation (2.8) as
\[ (\rho_{\gamma_j} e_j^\top)^{\ell}(\Gamma) = e_{j-1} \quad (j = 2, \ldots, n) \quad \text{and} \quad (\rho_{\gamma_1} e_1^\top)^{\ell}(\Gamma) = 0 \]
and then making use of the second formula in (2.8). Formulas (2.16) follow from (2.15) by the general principle (2.13).

2.3. Companion matrices. For a monic polynomial \( p \in \mathbb{F}[z] \), the associated left and right companion matrices are defined as
\[
C_L(p) = \begin{bmatrix}
0 & 0 & \cdots & 0 & -p_0 \\
1 & 0 & \cdots & 0 & -p_1 \\
0 & 1 & \cdots & 0 & -p_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & -p_{n-1}
\end{bmatrix} = C_R(p)^\top, \quad p(z) = z^n + \sum_{k=0}^{n-1} p_k z^k. \tag{2.18}
\]
By [17, Theorem A2], two matrices \( A, B \in \mathbb{F}^{n \times n} \) are similar over a unital ring \( \mathbb{F} \) if and only if the pencils \( zI_n - A \) and \( zI_n - B \) are equivalent over \( \mathbb{F}^{n \times n}[z] \) (i.e., one of them can be transformed into another by elementary row and column operations). Combining this result with the observation that the pencils \( zI_n - C_L(p) \) and \( zI_n - C_R(p) \) are both equivalent over \( \mathbb{F}^{n \times n}[z] \) to the diagonal polynomial matrix \( \begin{bmatrix} I_{n-1} & 0 \\
0 & p(z) \end{bmatrix} \), leads to the conclusion (see [17]) that
\[ C_L(p) \sim C_R(p) \quad \text{for any monic} \quad p \in \mathbb{F}[z]. \tag{2.19} \]
Remark 2.7. For a monic $p \in \mathbb{F}[z]$ and associated companion matrices $C_{\ell}(p)$ and $C_{r}(p)$ (as in (2.18)),

$$\mu_{C_{\ell}(p),r} = \mu_{C_{r}(p),r} = \Psi_{C_{\ell}(p),e_1} = \Psi_{e_1, C_{r}(p)} = p. \quad (2.20)$$

Indeed, recalling $e_j$, the $j$-th column of $I_n$ and observing that

$$C_{\ell}(p)^{k-1}e_1 = e_k \quad (k = 1, \ldots, n) \quad \text{and} \quad C_{\ell}(p)^ne_1 = -\sum_{j=1}^n e_jp_j, \quad (2.21)$$

we see that the minimal right linearly dependent set $\{C_{\ell}(p)^je_1\}_{j=1}^d$ occurs for $d = n$, and the relation

$$C_{\ell}(p)^ne_1 + \sum_{j=1}^n C_{\ell}(p)^{n-j}e_1p_{n-j} = 0$$

means that $\Psi_{C_{\ell}(p),e_1} = p$. Multiplying the latter equality by $C_{\ell}(p)^k$ on the left ($k = 1, \ldots, n - 1$) we get, on account of (2.21),

$$0 = C_{\ell}(p)^ne_k + \sum_{j=1}^n C_{\ell}(p)^{n-j}e_kp_{n-j} = \left(C_{\ell}(p)^n + \sum_{j=1}^n C_{\ell}(p)^{n-j}p_{n-j}\right)e_k = p^{\ell}(C_{\ell}(p))e_k,$$

where the second equality holds since $e_k \in Z^p_n$. Since $p^{\ell}(C_{\ell}(p))e_k = 0$ for $k = 1, \ldots, n$, it follows that $p^{\ell}(C_{\ell}(p)) = 0$, so that $p \in (\mu_{C_{\ell}(p),r})$. Since $p = \Psi_{C_{\ell}(p),e_1}$ is a divisor of $\mu_{C_{\ell}(p),r}$, it follows that $\mu_{C_{\ell}(p),r} = p$. The rest of (2.19) is verified similarly.

Remark 2.8. Although in general, $\mu_{C_{\ell}(p),r}$ and $\mu_{C_{r}(p),r}$ are not equal to $p$ (quite expectedly), they are divisible by $p$ on the right and on the left, respectively. The latter follows from the general principle (2.13) and equalities

$$\Psi_{e_1, C_{\ell}(p)} = \Psi_{C_{r}(p),e_n} = p$$

which are verified as in the proof of Proposition (2.6).

We next specify formulas (2.7) and (2.11) for the case where $A$ is a companion matrix (we will need them in Section 6). To this end, we recall the backward-shift operator $R_0$ acting on $\mathbb{F}[z]$ by the rule

$$R_0 : \sum_{j=0}^N f_jz^j \rightarrow \sum_{j=0}^{N-1} f_{j+1}z^j. \quad (2.22)$$

Proposition 2.9. Let $p \in \mathbb{F}[z]$ and $C_{\ell}(p)$ be defined as in (2.18). Given any $f \in \mathbb{F}[z]$, let us divide it by $p$ on the left:

$$f(z) = p(z)g(z) + b(z), \quad g \in \mathbb{F}[z], \quad b(z) = b_0 + b_1z + \ldots + b_{n-1}z^{n-1}, \quad (2.23)$$

and let us define the polynomials

$$g_j = (R_0^j p) \cdot g + R_0^j b \quad \text{for} \quad j = 1, \ldots, n. \quad (2.24)$$
Then $g_n = g$ and

$$(e_1 f)^{e_\epsilon}(C_\epsilon(p)) = b := \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix}, \quad L_{C_\epsilon(p)}(e_1 f) = G := \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g \end{bmatrix}. \quad (2.25)$$

Proof. Since $p$ is monic and $\deg p = n$, we have $R_0^n p = 1$. Since $\deg b < n$, we have $R_0^n b = 0$. Then for $j = n$, the formula (2.24) gives $g_n = g$. By definitions (2.24), we have

$$zg_{j+1}(z) + pjg(z) + b_{j+1} = z(R_0^{j+1} p)(z)g(z) + (R_0^j b(z) + pjg(z) + b_j$$

$$= [z(R_0^{j+1} p)(z) + pj]g(z) + (R_0^j b)(z) + b_j$$

$$= (R_0^j p)(z)g(z)g + R_0^j b = g_j(z)$$

for all $j = 1, \ldots, n - 1$. The latter $n - 1$ equalities along with (2.23) can be written in the matrix form as

$$\begin{bmatrix} f(z) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} z & 0 & \ldots & 0 & p_0 \\ -1 & z & \ldots & 0 & p_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & z + p_{n-1} \end{bmatrix} \begin{bmatrix} g_1(z) \\ g_2(z) \\ \vdots \\ g_n(z) \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix}.$$

Writing this identity in the form $e_1 f(z) = (zI_n - C_\epsilon(p))G(z) + b$, we arrive at both formulas in (2.25), by Remark 2.2.

The same arguments lead us to the right-sided analogues of formulas (2.25) which are recorded below for future references.

**Proposition 2.10.** Let $q \in \mathbb{F}[z]$ (deg $q = k$) and $C_\epsilon(p) \in \mathbb{F}^{k \times k}$ be defined as in (2.18). Given any $f \in \mathbb{F}[z]$, let us divide it by $q$ on the right

$$f(z) = h(z)q(z) + d(z), \quad h \in \mathbb{F}[z], \quad d(z) = d_0 + d_1 z + \ldots + d_{n-1} z^{k-1},$$

and let us define the polynomials

$$h_j = h \cdot (R_0^j q) + R_0^j d \quad \text{for} \quad j = 1, \ldots, n. \quad (2.26)$$

Then $h_k = h$ and

$$(fe_1^\top)^{Cr}(C_\epsilon(q)) = d := [d_0 \ldots d_{n-1}], \quad L_{C_r(q)}(fe_1^\top) = [h_1 \ldots h].$$

3. Controllable and Observable Pairs

For $A \in \mathbb{F}^{n \times n}$ and $v \in \mathbb{F}^{n \times 1}$, the input pair $(A, v)$ is called controllable, if its controllability matrix

$$C_{A, v} = \begin{bmatrix} v & Av & \ldots & A^{n-1}v \end{bmatrix}$$

is invertible. \quad (3.1)
In this case, the vector \( v \) is called a **right cyclic vector** for \( A \). For similar input pairs \((A, v)\) and \((A', v')\) (as in Definition 2.4), we have from (3.1)
\[
Tc_{A,v} = \begin{bmatrix} v' & TAT^{-1}v' & \ldots & TA^{n-1}T^{-1}v' \end{bmatrix} = c_{A',v'},
\]
from which we see that controllability is similarity-invariant.

The concept of controllability goes back to [10]; the discrete time-invariant linear system
\[
x(k + 1) = Ax(k) + vw(k), \quad x(0) = x_0
\]
is called controllable, if for any preassigned \( x(m) \in \mathbb{F}^n \), there exists the input sequence \( w(0), \ldots, w(m - 1) \in \mathbb{F} \) transferring the given initial state \( x(0) \) into \( x(m) \). The latter turns out to be equivalent to the controllability matrix (3.1) be invertible.

The concept of observability is dual to that of controllability: given \( A \in \mathbb{F}^{n \times n} \) and a row vector \( u \in \mathbb{F}^{1 \times n} \), the discrete time-invariant linear system
\[
x(k + 1) = Ax(k), \quad y(k) = vx(k), \quad x(0) = x_0
\]
is called observable if any preassigned output sequence \( y(0), \ldots, y(n - 1) \) can be generated by an appropriate initial state \( x(0) \). The latter holds if and only if the observability matrix of the pair \((u, A)\) is invertible which we adopt as the definition of observability:

Given \( A \in \mathbb{F}^{n \times n} \) and a row vector \( u \in \mathbb{F}^{1 \times n} \), the output pair \((u, A)\) is called observable if its observability matrix
\[
\Omega_{u,A} = \begin{bmatrix} u & \vdots & uA^{n-1} \\ uA & \vdots & uA^{n-1} \end{bmatrix}
\]
is invertible. (3.2)

In this case, \( u \) is called a **left cyclic vector** for \( A \). For similar output pairs \((u, A)\) and \((u', A')\), we have \( \Omega_{u',A'} = \Omega_{u,A}T^{-1} \) and hence, observability is similarity-invariant.

**Remark 3.1.** It follows from the definition (3.1) that an input pair \((A, v)\) with \( A \in \mathbb{F}^{n\times n} \) is controllable if and only if the least integer \( d \) for which (2.12) holds equals \( d = n \), i.e., if and only if \( \deg \mathfrak{V}_{A,v} = n \). Similarly, an output pair \((u, A)\) is observable if and only if \( \deg \mathfrak{F}_{u,A} = n \).

Explicit formulas for minimal polynomials of controllable and observable pairs are given below.

**Proposition 3.2.** (1) If the pair \((A, v)\) (with \( v \in \mathbb{F}^n \)) is controllable, then its minimal polynomial is given by the formula
\[
\mathfrak{P}_{A,v}(z) = z^n + \sum_{k=0}^{n-1} z^k b_k = z^n - \mathfrak{A}_{n}(z)\mathfrak{C}^{-1}_{A,v}A^n v, \quad (3.3)
\]
(where \( \mathfrak{A}_{n} \) is given by (1.1)), and furthermore,
\[
\mathfrak{C}^{-1}_{A,v}A\mathfrak{C}_{A,v} = C_{\ell}(\mathfrak{P}_{A,v}), \quad \mathfrak{C}^{-1}_{A,v}v = e_1,
\]
i.e., the pair \((A, v)\) is similar to the pair \((C_{\ell}(\mathfrak{P}_{A,v}), e_1)\).
(2) If the pair \((u, A)\) (with \(u \in \mathbb{F}^{1 \times n}\)) is observable, then
\[ \mathcal{P}_{u,A}(z) = z^n - uA^n \mathcal{D}_{u,A}^{-1} \mathcal{A}_n(z)^\top, \] (3.4)
and furthermore, the pair \((u, A)\) is similar to the pair \((e_1^\top, C_T(\mathcal{P}_{u,A}))\), as
\[ \mathcal{D}_{u,A}A\mathcal{D}_{u,A}^{-1} = C_T(\mathcal{P}_{u,A}) \quad \text{and} \quad u\mathcal{D}_{u,A} = e_1^\top. \]

Proof. If the pair \((A, v)\) is controllable, equality (2.12) holds with \(d = n\) and the coefficients \(b_k\) are defined by the formula
\[ \begin{bmatrix} b_0 \\ \vdots \\ b_{n-1} \end{bmatrix} = -c_{A,v}^{-1}A^n v. \] (3.5)
The formula (3.3) is now immediate. By (3.1), \(c_{A,v}^{-1} v = e_1\) and furthermore,
\[ c_{A,v}^{-1} A c_{A,v} = c_{A,v}^{-1} \begin{bmatrix} A v & \ldots & A^{n-1} v & A^n v \end{bmatrix} \]
\[ = \begin{bmatrix} e_2 & \ldots & e_{n-1} & c_{A,v}^{-1} A^n v \end{bmatrix} = C_e(\mathcal{P}_{A,v}) \] (3.6)
where \(C_e(\mathcal{P}_{A,v})\) is the left companion matrix of the polynomial (3.3). The second part is verified similarly. \(\square\)

By Remark 2.5, similar pairs have the same minimal polynomial. For controllable or observable pairs, we have the converse.

**Theorem 3.3.** Controllable pairs \((A, v)\) and \((A', v')\) (observable pairs \((u, A)\) and \((u', A')\)) are similar in the sense of Definition 2.4 if and only if their minimal polynomials \(\mathcal{P}_{A,v}\) and \(\mathcal{P}_{A',v'}\) (\(\mathcal{P}_{u,A}\) and \(\mathcal{P}_{u',A'}\)) are equal.

Proof. The “only if” part is contained in Remark 2.5. To justify the “only if” part, let us assume that the minimal polynomials \(\mathcal{P}_{A,v}\) and \(\mathcal{P}_{A',v'}\) of two controllable pairs \((A, v)\) and \((A', v')\) are equal. Then the matrices \(A\) and \(A'\) have the same dimensions (equal to \(\deg \mathcal{P}_{A,v} = \deg \mathcal{P}_{A',v'} = n\)).

The controllability matrices \(e_{A,v}\) and \(e_{A',v'}\) are both invertible, and we may let \(T := e_{A',v'} e_{A,v}^{-1}\). We next show that
\[ A'^j v' = T A^j v \quad \text{for} \quad j = 0, \ldots, n. \] (3.7)
Indeed, comparing the corresponding columns in the matrix equality
\[ T e_{A,v} = e_{A',v'} \]
gives equalities (3.7) for \(j = 0, \ldots, n - 1\). From the formula (3.5) for coefficients of \(\mathcal{P}_{A,v}\) and from the similar formula for coefficients of \(\mathcal{P}_{A',v'}\) we conclude (since \(\mathcal{P}_{A,v} = \mathcal{P}_{A',v'}\)) that \(e_{A,v}^{-1} A^n v = e_{A',v'}^{-1} A^n v'\), which can be written equivalently as \(A'^n v' = T A^n v\) thus justifying equality (3.7) for \(j = n\). Letting \(j = 0\) in (3.7) gives \(v' = T v\), while all other equalities in (3.7) imply
\[ A' e_{A', v'} = \begin{bmatrix} A' v' & \ldots & A^n v' \end{bmatrix} = \begin{bmatrix} T A v & \ldots & T A^n v \end{bmatrix} = T A e_{A,v}. \]
which is the same as \( A' = T A T^{-1} \), by the definition of \( T \). Thus, the pairs \((A, v)\) and \((A', v')\) are similar. The statement concerning observable pairs follows by similar arguments. \( \square \)

3.1. \( I_{A,v} \) and \( I_{u,A} \) as generic right and left ideals. As a consequence of Theorem 3.3, it follows that any right or left ideal in \( \mathbb{F}[z] \) is necessarily of the form (2.9), and that under controllability/observability assumption the representing pair \((A, v)\) or \((u, A)\) is unique up to similarity.

**Proposition 3.4.** (1) Any right ideal \( I \subset \mathbb{F}[z] \) is of the form

\[
I = I_{A,v} := \{ p \in \mathbb{F}[z] : (vp)^{\ell}(A) = 0 \} = \langle \mathfrak{P}_{A,v} \rangle_r
\]

for some controllable pair \((A, v)\). Controllable pairs \((A, v)\) and \((A', v')\) define the same ideal \( I_{A,v} = I_{A',v'} \) if and only if they are similar.

(2) Any left ideal \( I \subset \mathbb{F}[z] \) is of the form

\[
I = I_{u,A} := \{ p \in \mathbb{F}[z] : (pu)^{\ell}(A) = 0 \} = \langle \mathfrak{P}_{u,A} \rangle_{\ell}
\]

for some observable pair \((u, A)\). Observable pairs \((u, A)\) and \((u', A')\) define the same ideal \( I_{u,A} = I_{u',A'} \) if and only if they are similar.

**Proof.** Given a right ideal \( I \subset \mathbb{F}[z] \), let \( f \) be its generator. The pair \((C_{\ell}(f), e_1)\) is controllable and since \( u \mathfrak{P}_{C_{\ell}(f), e_1} = f \) (by (2.20)), we have

\[
I = \langle f \rangle_r = \langle \mathfrak{P}_{C_{\ell}(f), e_1} \rangle_r = I_{C_{\ell}(f), e_1}.
\]

By Theorem 3.3, two controllable pairs \((A, v)\) and \((A', v')\) are similar if and only if their minimal polynomials are equal, i.e.,

\[
I_{A,v} = \langle \mathfrak{P}_{A,v} \rangle_r = \langle \mathfrak{P}_{A',v'} \rangle_r = I_{A',v'}.
\]

This completes the proof of part (1). Part (2) follows similarly. \( \square \)

**Remark 3.5.** Any two-sided ideal \( I \subset \mathbb{F}[z] \) is of the form \( I = I_{A,v} \) for a unique (up to similarity) controllable pair \((A, v)\) such that \( C_{\ell} A_{A,v}^{1} A_{n} v \in \mathbb{Z}_{\mathbb{F}}^{n \times 1} \).

Alternatively, any two-sided ideal \( I \subset \mathbb{F}[z] \) is of the form \( I = I_{u,A} \) for a unique (up to similarity) observable pair \((u, A)\) such that \( u A_{n} A_{u,A}^{1} \in \mathbb{Z}_{\mathbb{F}}^{1 \times n} \).

By Propositions 3.2 and 3.4, the latter statement asserts that any two-sided ideal is a left or right ideal generated by a polynomial in \( Z_{\mathbb{F}}[z] \), which is obviously true.

We record several concrete examples of controllable and observable pairs that have already appeared above.

**Example 3.6.** For any monic \( p \in \mathbb{F}[z] \) of degree \( n \), the pairs \((C_{\ell}(p), e_1)\) and \((C_{r}(p), e_n)\) are controllable, while the pairs \((e_1, C_{r}(p))\) and \((e_n, C_{\ell}(p))\) are observable (since their minimal polynomials equal \( p \) and \( \deg p = n \)).

**Example 3.7.** If \( \Gamma_{\gamma} \in \mathbb{F}^{n \times n} \) is of the form (2.14), then the pair \((\Gamma_{\gamma}, e_1)\) is controllable, while the pair \((e_n^\top, \Gamma_{\gamma})\) is observable (since their minimal polynomials are of degree \( n \), by formulas (2.15) for \( k = 1 \) and \( k = n \), respectively).
For the next example, we recall the notion of polynomial independence (P-independence) introduced in [11]; see also [12, 13, 14].

**Definition 3.8.** A set \( \{\alpha_1, \ldots, \alpha_n\} \subset F \) is called left (right) P-independent if the monic linear polynomials \( \rho_{\alpha_1}, \ldots, \rho_{\alpha_n} \) are left (right) coprime.

**Proposition 3.9.** Let \( A \in F^{n \times n} \) be diagonal and let \( v = e_1 + \cdots + e_n \): 
\[
A = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \alpha_n \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. 
\]

The pair \((A, v)\) is controllable (the pair \((v^\top, A)\) is observable) if and only if the set \( \{\alpha_1, \ldots, \alpha_n\} \) is left (right) P-independent.

**Proof.** It follows from (2.2) and (2.7), that for every polynomial \( f \in F[z] \),
\[
(f^\ell)(A) = \begin{bmatrix} f^\ell(\alpha_1) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & f^\ell(\alpha_n) \end{bmatrix}, \quad (vf)^\ell(A) = \begin{bmatrix} f^\ell(\alpha_1) \\ \vdots \\ f^\ell(\alpha_n) \end{bmatrix},
\]
and hence, the ideals \( \mathbb{I}_{A, r} = \mathbb{I}_{A, v} \) consist of all polynomials that vanish on the left at \( \alpha_1, \ldots, \alpha_n \). Thus, \( \mathbb{I}_{A, r} = \mathbb{I}_{A, v} = \bigcap_{j=1}^n \langle \rho_{\alpha_j} \rangle_r \) and subsequently,
\[
\mu_{A, \ell} = \Psi_{A, v} = \text{lrcm}(\rho_{\alpha_1}, \ldots, \rho_{\alpha_n}).
\]

One can see from the definition (3.1) that the controllability matrix of the pair (3.8) is equal to the left Vandermonde matrix
\[
C_{A, v} = V_\ell(\alpha) := \begin{bmatrix} \alpha_1^{j-1} \\ \vdots \\ \alpha_n^{j-1} \end{bmatrix}_{i,j=1}^n, \quad \alpha := (\alpha_1, \ldots, \alpha_n).
\]

This matrix is invertible (i.e., the pair \((A, v)\) is controllable) if and only if \( \deg \Psi_{A, v} = n \) (we recall that \( \deg \Psi_{A, v} \) equals the maximal number of left-most right linearly independent columns in \( C_{A, v} \)). Due to (3.10), \( \deg \Psi_{A, v} = n \) if and only if the polynomials \( \rho_{\alpha_1}, \ldots, \rho_{\alpha_n} \) are left coprime, i.e., the set \( \{\alpha_1, \ldots, \alpha_n\} \) is left P-independent. The statement concerning the pair \((v^\top, A)\) is justified similarly. \( \square \)

The polynomials that are (left or right) minimal polynomials of an algebraic set in a division ring are called Wedderburn polynomials; we refer to [11, 12, 14, 15] for the thorough account on the subject. Since Wedderburn polynomials can be characterized as least common multiples of coprime monic linear polynomials, the formula (3.10) asserts that the pair \((A, v)\) is controllable if and only if its minimal polynomial \( \Psi_{A, v} \) is a Wedderburn polynomial.

In the next proposition we will use the minimal polynomial of a controllable pair to get a closed (and fairly explicit) formula for the least right common multiple of several given left-coprime polynomials.
Proposition 3.10. Given left-coprime monic $f_1, \ldots, f_k \in \mathbb{F}[z]$, let
\[
C_\ell(f_1, \ldots, f_k) := \begin{bmatrix} C_\ell(f_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_\ell(f_k) \end{bmatrix}, \quad E = \begin{bmatrix} e_1, n_1 \\ \vdots \\ e_1, n_k \end{bmatrix},
\]
where $n_j = \deg f_j$, and let $n := n_1 + \ldots + n_k$. Then the pair $(C_\ell(f_1, \ldots, f_k), E)$ is controllable and its minimal polynomial is given by
\[
\mathfrak{P}_{C_\ell(f_1, \ldots, f_k), E} = \operatorname{lrcm}(f_1, \ldots, f_k)
\]
(3.12)
\[
= z^n - \mathfrak{A}_n(z) C_\ell(f_1, \ldots, f_k), E \begin{bmatrix} C_\ell(f_1)^n e_1, n_1 \\ \vdots \\ C_\ell(f_k)^n e_1, n_k \end{bmatrix}.
\]

Proof. By the block-diagonal structure of $C_\ell(f_1, \ldots, f_k)$ and due to (2.20),
\[
\mathfrak{P}_{C_\ell(f_1, \ldots, f_k), E} = \operatorname{lrcm}(\mathfrak{P}_{C_\ell(f_1), e_1, n_1}, \ldots, \mathfrak{P}_{C_\ell(f_k), e_1, n_k}) = \operatorname{lrcm}(f_1, \ldots, f_k).
\]
Since $f_1, \ldots, f_k$ are left coprime,
\[
\deg(\operatorname{lrcm}(f_1, \ldots, f_k)) = \deg f_1 + \ldots + \deg f_k = n.
\]
Hence, the two last formulas imply $\deg \mathfrak{P}_{C_\ell(f_1, \ldots, f_k), E} = n$. Therefore, the matrix $C_\ell(f_1, \ldots, f_k), E$ is invertible (i.e., the pair $(C_\ell(f_1, \ldots, f_k), E)$ is controllable), and we can apply formula (3.3) to complete the proof of (3.12). \(\square\)

4. Cyclic matrices and similarity reduction

A matrix $A$ over a division ring $\mathbb{F}$ is called cyclic if it admits a (left or right) cyclic vector, i.e., if it can be embedded into a controllable or an observable pair. Alternatively, cyclic matrices can be defined as the ones similar to companion matrices or as the matrices having one non-constant invariant factor. All these equivalent definitions are recorded below.

Theorem 4.1. Given a matrix $A \in \mathbb{F}^{n \times n}$, the following are equivalent:

(1) There exists $v \in \mathbb{F}^{n \times 1}$ such that the pair $(A, v)$ is controllable.

(2) There exists $u \in \mathbb{F}^{1 \times n}$ such that the pair $(u, A)$ is observable.

(3) $A$ is similar to a (left or right) companion matrix.

(4) The pencil $zI_n - A$ is equivalent to a polynomial matrix $\begin{bmatrix} I_{n-1} & 0 \\ 0 & h(z) \end{bmatrix}$.

If this is the case (i.e., if $A$ is cyclic), then
(a) $A \sim C_\ell(\mathfrak{P}_{A, v}) \sim C_r(\mathfrak{P}_{u, A})$ for any cyclic vectors $v, u$ of $A$.

(b) If $A = TC_\ell(f)T^{-1} = SC_r(g)S^{-1}$ for some $f, g \in \mathbb{F}[z]$, then the vectors $v = Te_1$ and $u = e_1^T S^{-1}$ are cyclic for $A$ and furthermore,

\[
f = \mathfrak{P}_{A, v} \quad \text{and} \quad g = \mathfrak{P}_{u, A}.
\]

(c) The invariant factor $h$ of $A$ from part (4) is necessarily of the form $h = \mathfrak{P}_{A, w}$ for some cyclic vector $w$ of $A.$
Proof. Each one of the properties (1)-(4) is similarity invariant. By (3.6) and (3.2), \( A \) admits a right (left) cyclic vector if and only if it is similar to a left (right) companion matrix. By (2.19), we now conclude that the statements (1), (2), (3) are equivalent. Since (4) holds (with \( h = f \)) for any companion matrix \( C_\ell(f) \), the equivalence (3) \( \iff \) (4) follows.

The statement (a) follows from Proposition 3.2. If \( A = TC_\ell(f)T^{-1} \) and \( v = Te_1 \), then the input pairs \((A, v)\) and \((C_\ell(f), e_1)\) are similar (see Definition 2.4) and therefore, \( \Psi_{A, v} = \Psi_{C_\ell(f), e_1} = f \), by Remark 2.5 and due to (2.20). The rest of the part (b) follows from similarity of output pairs \((u, A)\) and \((e_1^T, C_\ell(g))\). Finally, being an invariant factor of \( A \), the polynomial \( h \) in (4) is also an invariant factor for its companion matrix \( C_\ell(h) \). Therefore, the pencils \( zI_n - A \) and \( zI_n - C_\ell(h) \) are equivalent and hence, \( A \sim C_\ell(h) \). Now part (c) follows from (b). \( \square \)

4.1. Similar polynomials. In the contrast to the commutative setting of Proposition 1.2, similar companion matrices over a noncommutative division ring do not have to be equal (for an example, take two similar elements \( \alpha \sim \alpha' \) (i.e., \( \alpha\beta = \beta\alpha' \) for some \( \beta \neq 0 \)) and consider the companion matrices \( C_\ell(p_\alpha) = \alpha \) and \( C_\ell(p_\alpha') = \alpha' \). The polynomials generating similar companion matrices are called similar; in notation: \( f \approx g \).

Proposition 4.2. For polynomials \( f, g \in \mathbb{F}[z] \), the following are equivalent:

1. \( \text{lrcm}(f, h) = hg \) for some \( h \in \mathbb{F}[z] \) such that \( f, h \) are left coprime.
2. \( \text{l lcm}(g, p) = fp \) for some \( p \in \mathbb{F}[z] \) such that \( p, g \) are right coprime.
3. \( fp = hg \) for some \( h, p \in \mathbb{F}[z] \) such that \( f, h \) are left coprime and \( p, g \) are right coprime.
4. \( C_\ell(f) \sim C_\ell(g) \).

Property (1) is the original definition of similar polynomials that appeared in [16]. The equivalence (1) \( \iff \) (2) was shown in [16, Theorem 1.18]. If \( f, g \) satisfy (1), then \( fp = hg \) for some \( p \in \mathbb{H}[z] \) which is necessarily right coprime with \( g \) (for otherwise, \( hg \) wouldn’t be the least right common multiple of \( f \) and \( h \)). On the other hand, if (3) is in force, then (1) holds with the same \( h \) (for otherwise, the polynomial \( \text{lrcm}(f, h) = f\tilde{p} = hg \) would be a proper left divisor of \( fp = hg \) implying that \( p = \tilde{p}q \) and \( g = gq \) for some non-constant \( q \in \mathbb{H}[z] \) contradicting the right coprimeness of \( p \) and \( g \)). Property (3) appears as the definition of polynomial similarity in [6, 8, 3] in terms of isomorphic cyclic modules. For the equivalence (3) \( \iff \) (4), see e.g., [15, Theorem 4.9].

As a consequence of Theorem 4.1, we have the following relaxed version of Theorem 3.3 (when similarity is imposed on state space matrices rather then on input or output pairs).

Proposition 4.3. (1) The minimal polynomials \( \Psi_{A, v} \) and \( \Psi_{A', v'} \) of controllable pairs \((A, v)\) and \((A', v')\) are similar if and only if \( A \sim A' \).
3.3, which is equivalent to relations (4.1).

Indeed, by Proposition 3.2, \( A \sim C_{\ell}(\mathcal{P}_{A,v}) \) and \( A' \sim C_{\ell}(\mathcal{P}_{A',v'}) \). By Theorem 4.1, we therefore have
\[
A \sim A' \iff C_{\ell}(\mathcal{P}_{A,v}) \sim C_{\ell}(\mathcal{P}_{A',v'}) \iff \mathcal{P}_{A,v} \approx \mathcal{P}_{A',v'}.
\]
The second statement follows similarly, due to (2.19).

Upon combining Theorem 4.1 and Theorem 3.3, we arrive at the following parametrization of the similarity class of a given polynomial.

**Theorem 4.4.** Given a monic \( p \in \mathbb{F}[z] \), the formula
\[
\varphi : v \mapsto \mathcal{P}_{C_{\ell}(p),v}
\]
establishes a map from the set of all cyclic vectors of the companion matrix \( C_{\ell}(p) \) onto the similarity class of \( p \). Moreover, \( \varphi(v) = \varphi(v') \) if and only if there exists an invertible \( T \in \mathbb{F}^{n \times n} \) such that
\[
TC_{\ell}(p) = C_{\ell}(p)T \quad \text{and} \quad v' =Tv.
\]

**Proof.** The vector \( e_1 \) is cyclic for \( C_{\ell}(p) \) and \( \mathcal{P}_{C_{\ell}(p),e_1} = p \). By Proposition 4.3 (part (1) with \( A = A' = C_{\ell}(p) \)),
\[
\mathcal{P}_{C_{\ell}(p),v} \approx \mathcal{P}_{C_{\ell}(p),e_1} = p
\]
for any cyclic vector \( v \) of \( C_{\ell}(p) \). Conversely, if \( g \approx p \), then \( C_{\ell}(p) \sim C_{\ell}(g) \), i.e., \( C_{\ell}(p) = T^{-1}C_{\ell}(g)T \) for some invertible \( T \in \mathbb{F}^{n \times n} \). If we let \( v = T^{-1}e_1 \), then controllable pairs \( (C_{\ell}(p),v) \) and \( (C_{\ell}(g),e_1) \) will be similar and hence,
\[
\mathcal{P}_{C_{\ell}(p),v} = \mathcal{P}_{C_{\ell}(g),e_1} = g,
\]
by Theorem 3.3. Therefore, the map \( \varphi \) is onto. Again due to Theorem 3.3, \( \mathcal{P}_{C_{\ell}(p),v} = \mathcal{P}_{C_{\ell}(p),v'} \) if and only if the pairs \( (C_{\ell}(p),v) \) and \( (C_{\ell}(p),v') \) are similar, which is equivalent to relations (4.1). \( \square \)

**4.2. Similarity reduction.** A cyclic matrix \( A \in \mathbb{F}^{n \times n} \) does not have to be similar to a two-diagonal matrix. To address Proposition 1.2 (part (4)) in the non-commutative setting, we put it in the following form: if the companion matrix \( C(f) \) \( (f \in \mathbb{C}[z]) \) is similar to a matrix \( \Gamma \) (1.3), then necessarily \( f = \rho_{\gamma_1} \cdots \rho_{\gamma_n} \). Two noncommutative extensions of the latter statement are given in Propositions 4.5 and 4.6 below.

**Proposition 4.5.** Let \( f \in \mathbb{F}[z] \) be a monic polynomial of degree \( n \). Let \( \Gamma_{\gamma} \), \( A \) and \( C_{\ell}(f_1, \ldots, f_k) \) be \( n \times n \) matrices given by (2.14), (3.8) and (3.11), respectively. Then
\[
\begin{align*}
(1) \quad & C_{\ell}(f) \sim \Gamma_{\gamma} \text{ if and only if } f \approx \rho_{\gamma_1} \rho_{\gamma_2} \cdots \rho_{\gamma_n}. \\
(2) \quad & C_{\ell}(f) \sim A = \text{diag}(\alpha_1, \ldots, \alpha_n) \text{ if and only if } f \approx \text{lrcm}(\rho_{\alpha_1}, \ldots, \rho_{\alpha_n}). \\
(3) \quad & C_{\ell}(f) \sim \text{diag}(C_{\ell}(f_1), \ldots, C_{\ell}(f_k)) \text{ if and only if } f \approx \text{lrcm}(f_1, \ldots, f_k).
\end{align*}
\]
All three statements are known in the more general setting of skew polynomials [15, Section 5]. In the present context, they follow from Proposition 4.3 and formulas (2.15), (3.10), (3.12) and (2.20).

The main point in part (1) is: $C_\ell(f)$ is similar to a two-diagonal matrix of the form (2.14) if and only if $f$ splits into the product of linear factors. Part (2) says that $C_\ell(f)$ is similar to a diagonal matrix if and only if $f$ is a Wedderburn polynomial. Since $\alpha_i = C_\ell(\rho_{\alpha_i})$, part (2) can be interpreted as the extremal particular case of part (3) when all diagonal blocks in the matrix $C_\ell(f_1, \ldots, f_k)$ are scalars. The opposite extremal case is the one where $f$ cannot be represented as the lrcm of its proper left divisors, or equivalently, the ideal $\langle f \rangle_r$ is irreducible in the sense that it is not contained into two distinct proper right ideals in $\mathbb{F}[z]$. Following Ore [16] we will call such polynomials indecomposable. By [16, Theorem 13, Part II], any polynomial $f \in \mathbb{F}[z]$ admits a representation

$$f = \text{lrcm}(f_1, \ldots, f_k)$$

where $f_1, \ldots, f_k$ are left coprime indecomposable polynomials, and this representation is unique up to similarity of each component.

We now present a more rigid version of Proposition 4.5 dealing with the fixed controllable pair $(C_\ell(f), e_1)$ rather than the companion matrix itself.

**Proposition 4.6.** Let $f \in \mathbb{F}[z]$ be a monic polynomial of degree $n$, let $\Gamma_\gamma$ and the controllable pairs $(A, v)$, $(C_\ell(f_1, \ldots, f_k), E)$ be defined as in (2.14), (3.8) and (3.11), respectively. Then

1. $(C_\ell(f), e_1) \sim (\Gamma_\gamma, e_1)$, i.e., there exists an invertible $T$ such that $TC_\ell(f)T^{-1} = \Gamma_\gamma$ and $Te_1 = e_1$

if and only if $f = \rho_{\gamma_1} \cdots \rho_{\gamma_n}$.

2. $(C_\ell(f), e_1) \sim (A, v)$ if and only if $f = \text{lrcm}(\rho_{\alpha_1}, \ldots, \rho_{\alpha_n})$.

3. $(C_\ell(f), e_1) \sim (C_\ell(f_1, \ldots, f_k), E)$ if and only if $f = \text{lrcm}(f_1, \ldots, f_k)$.

**Proof.** Since the pairs $(C_\ell(f), e_1)$, $(\Gamma_\gamma, e_1)$, $(A, v)$ and $(C_\ell(f_1, \ldots, f_k), E)$ are all controllable, any two of them are similar if and only if their minimal polynomials are equal. Since $\mathbb{F}_{C_\ell(f), e_1} = f$ the statements follow from the first formula in (2.15) (for $k = 1$) and formulas (3.10) and (3.12).

In case $\mathbb{F} = \mathbb{H}$, the skew field quaternions, some of the previous results can be elaborated a bit further, due to the facts that any non-real (i.e., non-central) element in $\mathbb{H}$ is algebraic of degree two and that $\mathbb{H}$ is algebraically closed on the left and on the right and hence any polynomial $f \in \mathbb{H}$ splits in $\mathbb{H}$. In this setting, any cyclic matrix is similar to a two-diagonal matrix (1.3), which is the exact analog of part (3) in Proposition 1.2. Part (3) in Proposition 4.5 is worked out to get the Jordan form of a cyclic matrix $A$ (which necessarily contains one block corresponding to each real eigenvalue and at most two blocks corresponding to each non-real eigenvalue), while part (3) in Proposition 4.6 eventually establishes similarity of a controllable pair $(A, v)$.
to the essentially unique pair \((\mathcal{J}, E)\) where \(\mathcal{J}\) is the block-diagonal matrix with diagonal blocks of the form (1.3) where this time, all \(\gamma_j\)'s are similar to each other and \(\gamma_{j+1} \neq \overline{\gamma}_j\) (the quaternion conjugate of \(\gamma_j\)). We omit details.

5. Ideal interpolation schemes

Characterizations of ideals of \(\mathbb{F}[z]\) given in Proposition 3.4 and Remark 3.5 in terms of evaluations (2.7) based on controllable and observable pairs suggest to take yet another look at interpolation problems in \(\mathbb{F}[z]\). We start with ideal interpolation schemes that were proposed in [1] in an attempt to come up with meaningful multivariate analogues of the Lagrange-Hermite interpolation problem. The single-variable non-commutative version of this concept is the following: given a ring \(\mathbb{F}\), a finite set \(\Phi = \{\phi_i\}_{i=1}^n\) of linearly independent functionals \(\phi_i: \mathbb{F}[z] \to \mathbb{F}\) is called a right (left, two-sided) ideal interpolation scheme if \(\bigcap_{i=1}^n \ker \phi_i\) is a right (left, two-sided) ideal in \(\mathbb{F}[z]\). Given an ideal interpolation scheme, the associated interpolation problem consists of finding all \(f \in \mathbb{F}[z]\) such that \(\phi_i(f) = c_i\) for preassigned \(c_j \in \mathbb{F}\) \((i = 1, \ldots, n)\). Since the problem is linear, the answer for a left (or right) scheme is given by the respective formulas (which are the same if the scheme is two-sided)

\[
f = f_0 + ph \quad \text{or} \quad f = f_0 + hp,
\]

(5.1)

where \(p\) is the generator of the ideal \(\bigcap_{i=1}^n \ker \phi_i\), where \(f_0\) is a unique particular solution to the problem with \(\deg f_0 < \deg p\), and where \(h \in \mathbb{F}[z]\) is a free parameter. The only remaining question is to find an explicit formula for \(f_0\) in terms of given \(c_1, \ldots, c_n\) and \(p\).

By Proposition 3.4, any right (left) ideal interpolation scheme in \(\mathbb{F}[z]\) can be embedded into the following left (right) interpolation problem with the interpolation condition given in terms of evaluations (2.7).

**Problem LP** \((A, v, b)\): given a controllable pair \((A, v)\) with \(A \in \mathbb{F}^{n \times n}\), and given \(b \in \mathbb{F}^{n \times 1}\), find a polynomial \(f \in \mathbb{F}[z]\) such that

\[
(vf)^{(e)}(A) = b.
\]

(5.2)

**Problem RP** \((B, u, d)\): given an observable pair \((u, B)\) with \(B \in \mathbb{F}^{k \times k}\), and given \(d \in \mathbb{F}^{1 \times k}\), find a polynomial \(f \in \mathbb{F}[z]\) such that

\[
(fu)^{(r)}(B) = d.
\]

(5.3)

The next two results specifying the parametrization formulas (5.1) in terms of interpolation data can be regarded as left and right noncommutative analogues of Proposition 1.1 (part (2)).

**Theorem 5.1.** The input pair \((A, v)\) with \(A \in \mathbb{F}^{n \times n}\) is controllable if and only if the problem **LP** \((A, v, b)\) has a solution for any \(b \in \mathbb{F}^{n \times 1}\). In this case, all solutions to the problem are parametrized by the formula

\[
f = f_\ell + \mathcal{P}_{A, v}h, \quad \text{with} \quad f_\ell(z) = \mathcal{A}_n(z)\mathcal{C}_{A, v}^{-1}b, \quad h \in \mathbb{F}[z],
\]

(5.4)
where $\mathfrak{A}_n$ is defined in (1.1), $\mathfrak{P}_{A,v}$ is the minimal polynomial of the pair $(A, v)$, $f_{\ell}$ is the low-degree solution, and $h$ is a free parameter.

**Proof.** by the division algorithm, the problem $LP(A, v, b)$ has a solution if and only if it has a low-degree one. To find a polynomial $f_{\ell}$ with $\deg f < n$ and subject to condition (5.2), we may take it in the form

$$f_{\ell}(z) = \sum_{j=0}^{n-1} f_j z^j = \mathfrak{A}_n(z)F, \quad F = \begin{bmatrix} f_0 \\ \vdots \\ f_{n-1} \end{bmatrix}, \quad (5.5)$$

and then compute, upon making use of (2.7) and (3.1),

$$(v f_{\ell})^e(A) = v f_0 + A v f_1 + \ldots + A^{n-1} v f_{n-1} = \mathfrak{C}_{A,v} F.$$

Thus the equation $(v f)^e(A) = \mathfrak{C}_{A,v} F = b$ has a solution $F$ for any $b \in F$ if and only if the controllability matrix $\mathfrak{C}_{A,v}$ is invertible, i.e., the pair $(A, v)$ is controllable. In this case, $f_{\ell}$ satisfies condition (5.2) if and only if $F = \mathfrak{C}^{-1}_{A,v} b$ which being substituted into (5.5), gives (5.4). Since the solution set of the homogeneous problem $LP(A, v, 0)$ is the right ideal $\langle \mathfrak{P}_{A,v} \rangle$, (by Proposition 3.4), the parametrization formula (5.4) follows. □

The right-sided version of Theorem 5.1 presented below is justified similarly.

**Theorem 5.2.** The output pair $(u, B)$ with $B \in F^{k \times k}$ is observable if and only if the problem $RP(B, u, d)$ has a solution for any $d \in F^{1 \times k}$. In this case, all solutions to the problem are given by the formula

$$f = f_r + h\mathfrak{P}_{u,B}, \quad \text{with} \quad f_r(z) = d \Sigma^{-1}_{u,B} \mathfrak{A}_k(z)^\top, \quad h \in F[z], \quad (5.6)$$

where $\mathfrak{P}_{u,A}$ is the minimal polynomial of the pair $(u, A)$, $f_r$ is the low-degree solution, and $h$ is a free parameter.

**Remark 5.3.** The problem (5.2) can be efficiently solved for any (not necessarily controllable) input pair $(A, v)$ as follows. Given a pair $(A, v)$, we find the smallest integer $d$ such that the vectors $v, Av, \ldots, A^{d-1}v$ are (right) linearly dependent and then construct the minimal polynomial $\mathfrak{P}_{A,v}$ (of degree $d$) as suggested in (2.12). The problem (5.2) has a solution if and only if the column $b$ belongs to the right range space of $\mathfrak{C}_{A,v}$ (i.e., to the right linear span of $v, Av, \ldots, A^{d-1}v$), the controllability space of the pair $(A, v)$. If this is the case, we represent $b$ as

$$b = \sum_{j=0}^{d-1} A^j vb_j, \quad \text{and let} \quad f_{\ell}(z) = \sum_{j=0}^{d-1} b_j z^j.$$

It is readily seen that all polynomials $f$ subject to the interpolation condition (5.2) are parametrized by the formula (5.4).

In conclusion we briefly address the first statement in Proposition 1.1. In the complex setting, $f(A)$ commutes with $A$, so the “only if” part is immediate. As complex polynomials respect similarity, the matrix $A$ can be taken in the canonical Jordan form, and then the commutativity relation
20 VLADIMIR BOLOTNIKOV

$AB = BA$ forces $B$ to be of triangular block Toeplitz structure. Then the Hermite-Lagrange polynomial with prescribed values (determined by $B$ at eigenvalues of $A$ (with multiplicities) satisfies $f(A) = B$.

In contrary to this case, polynomials over $\mathbb{F}$ do not respect similarity and besides, $f(A)$ does not have to commute with $A$. The solvability of the interpolation problem $f^\ell(A) = B$ for every $B \in \mathbb{F}^{n \times n}$ does not seem to have much to do with cyclicity of $A$. However, this problem falls in the left ideal interpolation scheme and its particular solution can be found recursively as follows. Letting $B = [b_1 \ b_2 \ \ldots \ b_n]$ we recall that $e_j \in \mathbb{Z}_F^n$ and write the interpolation condition $f^\ell(A) = B$ equivalently as

\[(e_j f)^\ell(A) = b_j \quad \text{for} \quad j = 1, \ldots, n. \quad (5.7)\]

Applying the procedure from Remark 5.2, we either conclude that the first condition in (5.7) is inconsistent (and hence the problem has no solutions) or we get all polynomials subject to this conditions in the form

\[ f = f^\ell + P_{A,e_1} h, \quad h \in \mathbb{F}[z]. \quad (5.8)\]

Making use of (2.8), we see that a polynomial $f$ of the form (5.8) satisfies conditions (5.2) for $j = 2, \ldots, n$ if and only if

\[ b_j = (e_j f)^\ell(A) = (e_j f^\ell)^\ell(A) + ((e_j P_{A,e_1})^\ell(A)) h^\ell(A) \]

for $j = 2, \ldots, n$, which can be written in terms of the parameter $h$ as

\[ (e_j h)^\ell(A) = d_j \quad \text{for} \quad j = 2, \ldots, n, \quad (5.9)\]

where $c_j = (e_j P_{A,e_1})^\ell(A)$ and $d_j = b_j - (e_j f^\ell)^\ell(A)$. Thus, either the problem (5.7) is inconsistent or it reduces (via (5.8)) to a similar problem (5.9) with fewer conditions. Continuing this reduction, we either conclude that the original problem (5.7) is inconsistent or will come up (in $n$ steps) with a parametrization of all its solutions.

6. Two-sided interpolation

Our next goal is to consider the problem which arises by combining left and right ideal interpolation schemes. We will call this problem two-sided as it contains both left and right interpolation conditions.

**Problem TSP:** Given a controllable pair $(A, v)$ and an observable pair $(u, B)$ (with $A \in \mathbb{F}^{n \times n}$, $B \in \mathbb{F}^{k \times k}$), along with the target vectors $b$ and $d$, find a polynomial $f \in \mathbb{F}[z]$ subject to conditions

\[(v f)^\ell(A) = b, \quad (f u)^r(B) = d, \quad \deg f < n + k. \quad (6.1)\]

By Theorems 5.1 and 5.2, the latter problem can be identified with the following one: given polynomials $p, q, f^\ell, f^r \in \mathbb{F}[z]$ with $\deg f^\ell < \deg p$ and $\deg f^r < \deg q$, find an $f \in \mathbb{F}[z]$ such that $\deg f < \deg p + \deg q$ and

\[ f - f^\ell \in \langle p \rangle^r \quad \text{and} \quad f - f^r \in \langle q \rangle^\ell. \quad (6.2)\]
Indeed, if we let
\[ p = \mathfrak{P}_{A,v} \quad \text{and} \quad q = \mathfrak{P}_{u,B}, \]
be minimal polynomials of the pairs \((A, v)\) and \((u, B)\), and then let
\[ f\ell = \mathfrak{A}_n c_{A,v}^{-1} b \quad \text{and} \quad f_r = d\Omega_{A,v}^{-1} A_{k}, \]
then conditions (6.2) turn out to be identical to parametrization formulas (5.4), (5.6), and hence, they are equivalent to conditions (5.2) and (5.3).

**Remark 6.1.** The degree constraint in (6.1) is not restrictive. By the left and right division algorithms, any polynomial \( f \in \mathbb{F}[z] \) can be uniquely represented as
\[ f = \tilde{f} + phq \quad \text{for some} \quad \tilde{f}, h \in \mathbb{F}[z], \quad \deg \tilde{f} < \deg p + \deg q, \]
and furthermore, \( f \) satisfies conditions (6.2) if and only if \( \tilde{f} \) does.

Theorem 6.11 below states that the problem (6.1) has a solution if and only if the Sylvester equation \( AY - YB = bu - vd \) has a solution \( Y \in \mathbb{F}^{n \times k} \).

The “only if” part of this criterion is verified in the next section via certain “two-sided” evaluation calculus.

6.1. **Two-sided evaluation.** Left and right evaluations (2.7) based on input and output pairs evaluate a scalar polynomial \( f \in \mathbb{F}[z] \) at these pairs, rather at matrices (as in (2.2)). In formula (6.6) below, we introduce a map \( \mathbb{F}[z] \to \mathbb{F}^{n \times k} \) that evaluates \( f \) at the couple \( \{(A, v), (u, B)\} \) consisting of an input pair \((A, v)\) and an output pair \((u, B)\).

Let us extend the backward-shift operators \( L_A \) and \( R_B \) defined via formulas (2.11) on vector polynomials of special form (polynomial multiples of a constant vector) to matrix polynomials of the form
\[ F = vf u, \quad \text{where} \quad v \in \mathbb{F}^{n \times 1}, \quad u \in \mathbb{F}^{1 \times k}, \quad f(z) = \sum f_j z^j \in \mathbb{F}[z], \]
by the formulas
\[ L_A(vf u) := L_A(vf) \cdot u = \sum_{j+k=0}^{\deg f-1} A^j vf_{k+j+1} u z^k, \quad (6.5) \]
\[ R_B(vf u) := v \cdot R_B(fu) = \sum_{j+k=0}^{\deg f-1} vf_{k+j+1} u B^j z^k. \]

Evaluating the top formula at \( B \) on the right and the bottom formula at \( A \) on the left we get the same outcomes which we will refer to as the **two-sided evaluation** of \( f \) at the couple \( \{(A, v), (u, B)\} \):
\[ (vf u)^{\ell\ell}(A, B) := \sum_{i+j=0}^{\deg f-1} A^i vf_{i+j+1} u B^j \]
\[ = (L_A(ufv))^\ell\ell(B) = (R_B(ufv))^\ell\ell(A). \]

\[ (vf u)^{\ell\ell}(A, B) := \sum_{i+j=0}^{\deg f-1} A^i vf_{i+j+1} u B^j \]
\[ = (L_A(ufv))^\ell\ell(B) = (R_B(ufv))^\ell\ell(A). \]
Proposition 6.2. For any \( f \in \mathbb{F}[z] \), the evaluations (2.7) and (6.6) satisfy the Sylvester equation

\[
A \cdot (vf u)^{\ell s}(A, B) - (vf u)^{\ell s}(A, B) \cdot B = (vf)^{\ell s}(A) \cdot u - v \cdot (fu)^{\ell r}(B). \tag{6.7}
\]

Indeed, by making use of polynomial expressions in (2.7) and (6.7) we get

\[
A \cdot (vf u)^{\ell s}(A, B) - (vf u)^{\ell s}(A, B) \cdot B = \sum_{i+j=0}^{\deg f} A^i v_{i+j+1} u B^j - \sum_{i+j=0}^{\deg f} A^i v_{i+j+1} u B^{j+1}
\]

\[
\sum_{j=1}^{\deg f} A^j v_j u B^j = (vf)^{\ell s}(A) \cdot u - v \cdot (fu)^{\ell r}(B).
\]

Corollary 6.3. If a polynomial \( f \) satisfies conditions (6.1), then the matrix \( Y = (vf u)^{\ell s}(A, B) \) solves the Sylvester equation \( AY - YB = bu - vd \).

For a concrete example, we will compute the two-sided evaluation at \( ((C_{\ell}(p), e_{1,n}), (e_{1,k}, C_{r}(q))) \) where \( C_{\ell}(p) \) and \( C_{r}(q) \) are companion matrices of given polynomials \( p, q \in \mathbb{F}[z] \) with \( \deg p = n, \deg q = k \). We let for short,

\[
Y^f = [\gamma_{ij}] := (e_{1,n}f e_{1,k})^{\ell s}(C_{\ell}(p), C_{r}(q)) \tag{6.8}
\]

and denote by \( \gamma_j \) and \( \gamma_i \) the \( j \)-th column and the \( i \)-th row of the matrix \( Y^f \). Upon specifying two last formulas in (6.6) to the present setting and combining Propositions 2.9 and 2.10 we see that

\[
\gamma_j = (e_{1,n}h_j)^{\ell s}(C_{\ell}(p)) \quad \text{and} \quad \gamma_i = (g_i e_{1,k})^{\ell r}(C_{r}(q)), \tag{6.9}
\]

where the polynomials \( g_i \) and \( h_j \) are defined in (2.24) and (2.26), respectively.

Remark 6.4. \( Y^f = 0 \) if and only if \( f = \alpha + phq \) for some \( h \in \mathbb{F}[z] \) and \( \alpha \in \mathbb{F} \).

Proof. Since \( q \) is the minimal polynomial of the pair \( (e_{1,k}, C_{r}(q)) \), it follows from (6.9) that \( \gamma_i = 0 \) if and only if \( g_i \) is a left multiple of \( q \). Given \( f = \alpha + phq \), equalities (2.23) hold with \( b(z) = b_0 = \alpha \) and \( g = hq \). Hence, \( g_i = (R_i^p) \cdot g \) is a left multiple of \( q \). Therefore, \( \gamma_i = 0 \) for \( i = 1, \ldots, k \) and hence, \( \gamma_i = 0 \). Conversely, if \( \gamma_i = 0 \), then \( g_j = r_j q \) for \( j = 1, \ldots, n \). In particular, \( g_n = g = r_n q \), and now it follows from (2.24) that \( R_i^b 0 = 0 \) for \( j = 1, \ldots, n - 1 \). Since \( \deg b < n \), the latter equalities imply that \( b = b_0 \). Then (2.23) takes the form \( f = pr_n q + b_0 \) as desired. \( \square \)

Since any \( f \in \mathbb{F}[z] \) can be represented as in (6.4) and then necessarily \( Y^f = Y_i^f \), by Remark 6.4, it follows that it suffices to compute \( Y^f \) for \( f \) with \( \deg f < n + k \). Any such polynomial can be represented as

\[
f = pg + b = hq + v \tag{6.10}
\]
with $b, h$ of degree less than $n$ and $d, g$ of degree less than $k$, i.e.,

\[
\begin{align*}
b(z) &= b_0 + \ldots + b_{n-1}z^{n-1}, \quad g(z) = g_0 + \ldots + g_{k-1}z^{k-1}, \\
d(z) &= d_0 + \ldots + d_{k-1}z^{k-1}, \quad h(z) = h_0 + \ldots + h_{n-1}z^{n-1}.
\end{align*}
\]

(6.11)

Below, we compute the matrix $\Upsilon^f$ in terms of polynomials (6.11).

**Lemma 6.5.** If $f$ is of the form (6.10), (6.11), then the columns $\Upsilon_j$ and the rows $\tilde{\Upsilon}_i$ of the matrix (6.8) are given by

\[
\Upsilon_k = \begin{bmatrix}
h_0 \\
\vdots \\
h_{n-1}
\end{bmatrix}, \quad \tilde{\Upsilon}_n = \begin{bmatrix}
g_0 & \ldots & g_{k-1}
\end{bmatrix},
\]

(6.12)

\[
\Upsilon_j = \sum_{i=0}^{k-j} C_{\ell}(p)^i \Upsilon_{kq_{j+i}} + \sum_{i=0}^{k-j-1} C_{\ell}(p)^i e_1d_{j+i} \quad (j = 1, \ldots, k-1),
\]

(6.13)

\[
\tilde{\Upsilon}_i = \sum_{j=0}^{k-i} p_{i+j} \Upsilon_n C_\ell(q)^j + \sum_{j=0}^{n-i-1} b_{i+j} e_1^T C_\ell(q)^j \quad (i = 1, \ldots, n-1)
\]

(6.14)

**Proof.** By (6.9) and (2.26) (for $j = k$),

\[
\Upsilon_k = (e_1h)^{\ell}(C_\ell(p))
\]

and since \(\deg h < \deg p\), the latter column consists of the coefficients of $h$, by Proposition 2.9, which verifies the first formula in (6.12). We next compute

\[
\Upsilon_j = (e_1h)^{\ell}(C_\ell(p))
\]

\[
= (e_1(h \cdot (R_0^j q) + R_0^j d))^{\ell}(C_\ell(p))
\]

\[
= (e_1(h \cdot (R_0^j g))^{\ell}(C_\ell(p)) + (e_1R_0^j d)^{\ell}(C_\ell(p))
\]

\[
= ((e_1h)^{\ell}(C_\ell(p)) \cdot R_0^j g)^{\ell}(C_\ell(p)) + (e_1(R_0^j d))^{\ell}(C_\ell(p))
\]

\[
= (\Upsilon_k \cdot R_0^j g)^{\ell}(C_\ell(p)) + (e_1(R_0^j d))^{\ell}(C_\ell(p))
\]

where we used (6.9) and (2.26) for the two first steps, the additivity of evaluation operators and the multiplicative property (2.8) (applied to $h$ and $R_0^j g$) for the next two steps, and finally we used (6.14) for the last step. The expression on the right side is the same as in (6.13), by definition (2.7) of left evaluation. Equalities for the rows $\tilde{\Upsilon}_i$ are verified similarly. \qed

6.2. **Sylvester equations.** We now consider the Sylvester equation

\[
C_\ell(p)X - XC_\ell(q) = be_{1,k}^T - e_{1,n}d
\]

(6.15)
with given $b = [b_0 \ldots b_{n-1}]^\top$ and $d = [d_0 \ldots d_{k-1}]$ and unknown $X \in \mathbb{F}^{n \times k}$, along with two associated “generalized” Sylvester equations

$$(xq)^{\epsilon_l}(C_{0}(p)) := \sum_{i=0}^{k} C_{0}(p)^{i}xq_i = b - \sum_{i=0}^{k-1} C_{0}(p)^{i}e_{1,n}d_i, \quad (6.16)$$

$$(p\bar{x})^{\epsilon_r}(C_{r}(q)) := \sum_{j=0}^{n} p_j\bar{x}C_{r}(q)^{j} = d - \sum_{j=0}^{n-1} b_je_{1}^\top C_{r}(q)^{j}. \quad (6.17)$$

with unknowns $x \in \mathbb{F}^{n \times n}$ and $\bar{x} \in \mathbb{F}^{1 \times k}$. The next result establishes one-to-one correspondences between four sets: solution sets of equations (6.15), (6.16), (6.17), and the set of all polynomials $f$ such that

$$(e_{1,n}f)^{\epsilon_l}(C_{0}(p)) = b, \quad (fe_{1,k}^\top)^{\epsilon_r}(C_{r}(q)) = d, \quad \Upsilon f = X, \quad \deg f < n + k, \quad (6.18)$$

where $\Upsilon f$ is defined in (6.8).

**Theorem 6.6.** Let $X_j$ and $\bar{X}_i$ denote the $j$-th column and the $i$-th row of a matrix $X \in \mathbb{F}^{n \times k}$. Then

1. If $f \in \mathbb{F}[z]$ satisfies conditions (6.18), then $X = \Upsilon f$ is a solution to the equation (6.15).

2. If $X$ solves the equation (6.15), then

   a. $x = X_k$ solves the equation (6.16);

   b. $\bar{x} = \bar{X}_n$ solves the equation (6.17);

   c. The formulas (we recall (1.1))

   $$f(z) = d\mathbb{A}_k(z)^\top + \mathbb{A}_n(z)Xe_{k,k}q(z)$$

   $$= \mathbb{A}_n(z)b + p(z)e_{n,n}^\top X\mathbb{A}_k(z)^\top \quad (6.19)$$

   define a unique $f \in \mathbb{F}[z]$ subject to conditions (6.18).

3. If $x \in \mathbb{F}^{n \times 1}$ solves the equation (6.16), then the formula

   $$f(z) = d\mathbb{A}_k(z)^\top + \mathbb{A}_n(z)xq(z) \quad (6.20)$$

defines a unique $f \in \mathbb{F}[z]$ subject to conditions (6.18), whereas the columns $X_k = x$ and

$$X_j = \sum_{i=0}^{k-j} C_{e}(p)^{i}xq_{j+i} + \sum_{i=0}^{k-j-1} C_{e}(p)^{i}e_{1,n}d_{j+i} \quad (1 \leq j \leq k - 1), \quad (6.21)$$

define the only solution $X \in \mathbb{F}^{n \times k}$ to equation (6.15) with $X_k = x$.

4. If $\bar{x}$ solves the equation (6.17), then the formula

   $$f(z) = \mathbb{A}_n(z)b + p(z)\bar{x}\mathbb{A}_k(z)^\top \quad (6.22)$$
defines a unique \( f \in \mathbb{F}[z] \) subject to conditions (6.18), whereas the rows 
\[ \bar{X}_n = \bar{x} \]
define the only solution \( X \in \mathbb{F}^{n \times k} \) to the equation (6.15) with \( \bar{X}_n = \bar{x} \).

**Proof.** Part (1) follows from Corollary 6.3 specialized to the present setting. Making use of the explicit formula
\[
C_r(q) = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1 \\
-q_0 & -q_1 & -q_2 & \ldots & -q_k-1
\end{bmatrix},
\]
we equate the corresponding columns in (6.15):
\[
C_\ell(p)X_1 + X_k q_0 = b - e_1, n d_0, \quad (6.23)
\]
\[
C_\ell(p)X_j - X_{j-1} + X_k q_{j-1} = -e_1, n d_{j-1} \quad \text{for} \quad j = 2, \ldots, k. \quad (6.24)
\]
From (6.24) we recursively recover \( X_{k-1}, X_{k-2}, \ldots, X_1 \) from \( X_k \) arriving at formulas (6.21) which are the same as (6.13). Substituting the formula (6.21) (for \( j = 1 \)) into (6.23), and moving all terms not containing \( X_k \) to the right side, we get the equality
\[
\sum_{i=0}^{k} C_\ell(p)^i X_k q_i = b - \sum_{i=0}^{k-1} C_\ell(p)^i e_1, n d_i, \quad (6.25)
\]
which means that \( x = X_k \) is a solution to the equation (6.16). As is easily seen, the system of equalities (6.25), (6.21) is equivalent to the system (6.23), (6.24) (i.e., to the Sylvester equality (6.15)). Therefore, with the fixed column \( X_k = x \) subject to (6.25), the only way to extend it to a solution \( X \) to the equation (6.15), is to use recursive formulas (6.21). This completes the proof part (2a) and the second half of part (3).

To prove the first half, take any \( x \in \mathbb{F}^{n \times 1} \) subject to (6.16) and extend it to the matrix \( X \) subject to (6.15) using formulas (6.21). Since the formulas (6.21) are the same as in (6.13), the matrix \( X \) equals to the matrix \( \Upsilon f \) corresponding to the polynomial
\[
f = h q + d, \quad \text{where} \quad h(z) = A_n(z) \cdot x \quad \text{and} \quad d(z) = d \cdot A_k(z)^\top, \quad (6.26)
\]
which has been announced in (6.20). Thus, this \( f \) satisfies the third condition in (6.18), and it is follows from (6.26) (by Proposition 2.10) that it also satisfies the second one, and that \( \deg f \leq \deg h + \deg q < n + k \). The two last conditions in (6.18) fix the quotient \( h \) and the remainder \( d \) of \( f \) (of degree less than \( n + k \)) when divided by \( q \) on the right and hence, determine
To show that $f$ of the form (6.26) also satisfies the first condition in (6.18), let us observe that $x$ can be interpreted as the left value $x = (e_1, n)^\ell (C_{\ell}(p))$.

Making use of the rule (2.25), we now can write (6.16) as

$$(xq)^\ell (C_{\ell}(p)) = (e_1, n)^\ell (C_{\ell}(p)) = b - (e_1, n)^\ell (C_{\ell}(p)),$$

which in turn, is equivalent (due to (6.26)) to

$$b = (e_1, n)^\ell (C_{\ell}(p)) + (e_1, n)^\ell (C_{\ell}(p)) = (e_1, n)^\ell (C_{\ell}(p)),$$

which completes the proof of part (3). Parts (4) and (2b) are verified similarly. It remains to confirm the part (2a). To this end, take any $X \in \mathbb{F}^{n \times k}$ subject to (6.15) and observe that

$$x := X_k = X_{e_{n,n}}$$

solve the respective equations (6.16) and (6.17) (by parts (1a) and (1b)) and hence the formulas (6.20) and (6.22) define a unique (and therefore, the same) polynomial $f$ subject to conditions (6.18). The formulas in (6.19) follow from (6.20), (6.22) and (6.27).

### 6.3. Quasi-ideals in $\mathbb{F}[z]$.

An additive subgroup $\mathcal{Q}$ of an associative ring $\mathcal{A}$ such that $\mathcal{Q} \mathcal{A} \cap \mathcal{A} \mathcal{Q} \subseteq \mathcal{Q}$ (called a quasi-ideal in [18]) amounts, in the setting of $\mathbb{F}[z]$, to the intersection of a left and a right ideal

$$\mathcal{Q}_{p,q} := \langle p \rangle_\ell \cap \langle q \rangle_\ell$$

generated by two given polynomials. Any element $f \in \mathcal{Q}_{p,q}$ is characterized by factorizations $f = pq = hq$ or by homogeneous interpolation conditions $(e_1, n)^\ell (C_{\ell}(p)) = 0$ and $(fe_1, k)^\ell (C_{\ell}(q)) = 0$. By letting $b = 0$ and $d = 0$ throughout Section 6.2 we arrive at the following result.

**Proposition 6.7.** Given polynomials $p, q \in \mathbb{F}[z]$, the formula

$$X \mapsto \mathfrak{A}_n(z) X e_{k, q}(z) = p(z) e_{n, n}^T X \mathfrak{A}(z)^T$$

establishes the one-to-one correspondence between the matrices $X \in \mathbb{F}^{n \times k}$ such that $C_{\ell}(p) X = X C_{\ell}(q)$ and the polynomials $f \in \mathcal{Q}_{p,q}$ of degree $\deg f < \deg p + \deg q$.

By Remark 6.1, any $f \in \mathcal{Q}_{p,q}$ can be uniquely represented as in (6.4) with $\tilde{f} \in \mathcal{Q}_{p,q}$; hence the description of the whole $\mathcal{Q}_{p,q}$ follows from Proposition 6.7.
6.4. Two-sided interpolation problems. As an intermediate step toward solving the problem TSP (6.1), we will consider the augmented two-sided problem ATSP (equivalent to the problem (6.18)) whose data set
\[
\Omega = \{(A, v), (u, B), b, d, S\} \tag{6.28}
\]
contains, a controllable pair \((A, v)\), an observable pair \((u, B)\) and the target values \(b \in \mathbb{F}^{n \times 1}\), \(d \in \mathbb{F}^{1 \times k}\) and \(S \in \mathbb{F}^{n \times k}\) for left, right and two-sided interpolation conditions. The formal definition of the problem is as follows.

**Problem ATSP**(\(\Omega\)): Given \(\Omega\) as in (6.28), find all \(f \in \mathbb{F}[z]\) subject to conditions (6.1) and
\[
(vfu)^{ts}(A, B) = S. \tag{6.29}
\]

By Proposition 3.2, the pairs \((A, v)\) and \((u, B)\) are similar to \((C_\ell(p), e_{1,n})\) and \((e_{1,k}^\top, C_r(q))\), respectively; in more detail,
\[
C_{A,v}^{-1}A\mathcal{C}_{A,v} = C_\ell(p), \quad C_{A,u}^{-1}v = e_1, \quad \mathcal{D}_{u,B}B\mathcal{D}_{u,B}^{-1} = C_r(q), \quad u\mathcal{D}_{u,B}^{-1} = e_1^\top. \tag{6.30}
\]

Since similar pairs have the same minimal polynomials, we will use notation (6.3) (i.e., \(p = \mathfrak{P}_{A,v}\) and \(q = \mathfrak{P}_{u,B}\)) throughout this section.

**Remark 6.8.** Conditions (6.1) and (6.29) can be equivalently written as
\[
(e_{1,n}f)^{\ell e}(C_\ell(p)) = b' := C_{A,v}^{-1}b, \tag{6.31}
\]
\[
(fe_{1,k}^\top)^{re}(C_r(q)) = d' := d\mathcal{D}_{u,B}^{-1}, \tag{6.32}
\]
\[
(e_{1,n}fe_{1,k}^\top)^{ts}(C_\ell(p), C_r(q)) = S' := C_{A,v}^{-1}S\mathcal{D}_{u,B}^{-1}. \tag{6.33}
\]

Indeed, due to equalities (6.30), we have for any \(f \in \mathbb{F}[z]\),
\[
(e_{1,n}f)^{\ell e}(C_\ell(p)) = C_{A,v}^{-1} \cdot (vf)^{\ell e}(A),
\]
\[
(fe_{1,k}^\top)^{re}(C_r(q)) = (fu)^{re}(B) \cdot \mathcal{D}_{u,B}^{-1},
\]
by Remark 2.5, while for the two-sided evaluation, we have from (6.6),
\[
(e_{1,n}fe_{1,k}^\top)^{ts}(C_\ell(p), C_r(q)) = C_{A,v}^{-1} \cdot (vf)^{ts}(A, B) \cdot \mathcal{D}_{u,B}^{-1}.
\]

Now we see from the latter equalities that interpolation conditions (6.31)–(6.33) are obtained from (6.1), (6.29) upon multiplying the latter by invertible \(T\) and \(\tilde{T}\) on the left and/or on the right respectively, and hence, the asserted equivalence follows.

As we know from Theorem 6.6, the interpolation problem (6.31)–(6.33) has a solution if and only if
\[
C_\ell(p)S' - S'C_r(q) = b'e_{1,k}^\top - e_{1,n}d', \tag{6.34}
\]
in which case the only solution is given by (6.19), i.e.,
\[
f = \mathfrak{A}_n b' + pe_{n,n}^\top S'\mathfrak{A}_k^\top = d'\mathfrak{A}_k^\top + \mathfrak{A}_n S' e_{k,k}q. \tag{6.35}
\]

Using relations (6.30) and the rightmost equalities in (6.31)–(6.33), we may write (6.34) and (6.35) in terms of \(\Omega\) to arrive at the following result.
Theorem 6.9. The problem ATSP has a solution if and only if
\[ \mathbf{A} \mathbf{S} - \mathbf{B} \mathbf{u} = \mathbf{b} \mathbf{u} - \mathbf{v} \mathbf{d}, \]
(6.36)
in which case the only solution is given by either formula
\begin{align*}
\mathbf{f}(z) &= \mathbf{A}_n(z) \mathbf{C}_{\mathbf{A},n}^{-1} \mathbf{b} + \mathbf{P}_{\mathbf{A},v}(z) \mathbf{e}_{n,n}^\top \mathbf{C}_{\mathbf{A},v}^{-1} \mathbf{S} \mathbf{D}_{\mathbf{u},B}^{-1} \mathbf{A}_k(z)^\top \\
&= \mathbf{d} \mathbf{D}_{\mathbf{u},B}^{-1} \mathbf{A}_k(z)^\top + \mathbf{A}_n(z) \mathbf{C}_{\mathbf{A},v}^{-1} \mathbf{S} \mathbf{D}_{\mathbf{u},B}^{-1} \mathbf{e}_{k,k} \mathbf{P}_{\mathbf{u},B}. \tag{6.37}
\end{align*}

Note that since the first terms in formulas (6.37) solve the respective one-sided problems (5.2) and (5.3) and the second terms are multiples of \( \mathbf{P}_{\mathbf{A},v} \) and \( \mathbf{P}_{\mathbf{u},B} \), respectively, it is immediate that \( \mathbf{f} \) defined in (6.37) satisfies conditions (6.1). A nontrivial part here is that \( \mathbf{f} \) also satisfies the third condition (6.29) and that two formulas in (6.37) represent the same polynomial. In particular, it follows from (6.37) that the problem ATSP(\( \Omega \)) is redundant: if (6.36) holds and \( \mathbf{f} \) satisfies (6.29) and any one of the two conditions in (6.1), then it also satisfies the second. In fact, the condition (6.29) alone determines the polynomial \( \mathbf{f} \) of degree less than \( n + k \) up to a constant.

Theorem 6.10. Given \( \Omega = \{(\mathbf{A}, \mathbf{v}), (\mathbf{u}, \mathbf{B}), \mathbf{S}\} \) as above, there is a polynomial \( \mathbf{f} \in \mathbb{F}[z] \) subject to condition (6.33) if and only if
\[ (\mathbf{I}_n - \mathbf{e}_{1,n} \mathbf{e}_{1,n}^\top) \mathbf{C}_{\mathbf{A},v}^{-1} (\mathbf{A} \mathbf{S} - \mathbf{B} \mathbf{u}) \mathbf{D}_{\mathbf{u},B}^{-1} (\mathbf{I}_k - \mathbf{e}_{1,k} \mathbf{e}_{1,k}^\top) = 0, \]
(6.38)
in which case a solution is uniquely defined (up to an arbitrary additive constant \( \alpha \in \mathbb{F} \)) by the formula
\begin{align*}
\mathbf{f}(z) &= \alpha + \mathbf{A}_n(z) \mathbf{C}_{\mathbf{A},v}^{-1} (\mathbf{A} \mathbf{S} - \mathbf{B} \mathbf{u}) \mathbf{D}_{\mathbf{u},B}^{-1} \mathbf{e}_{1,n} \\
&\quad + \mathbf{P}_{\mathbf{A},v}(z) \mathbf{e}_{n,n}^\top \mathbf{C}_{\mathbf{A},v}^{-1} \mathbf{S} \mathbf{D}_{\mathbf{u},B}^{-1} \mathbf{A}_k(z)^\top. \tag{6.39}
\end{align*}

Proof. As in the proof of the previous theorem, we pass to the equivalent interpolation problem (6.33) (with \( p \) and \( q \) as in (6.3)). If there is a polynomial \( \mathbf{f} \in \mathbb{F}[z] \) satisfying (6.33), then equality (6.34) holds for some \( \mathbf{b}' \) and \( \mathbf{d}' \). Then we also have
\[ (\mathbf{I}_n - \mathbf{e}_{1,n} \mathbf{e}_{1,n}^\top)(\mathbf{C}_\ell(p) \mathbf{S}' - \mathbf{S}' \mathbf{C}_r(q)) (\mathbf{I}_k - \mathbf{e}_{1,k} \mathbf{e}_{1,k}^\top) = 0, \]
(6.40)
which is the same as (6.38), due to relations (6.30) and the rightmost equality in (6.33). Conversely, if (6.38) holds, we see from (6.40) (which is equivalent to (6.38)) that equality (6.34) holds for
\[ \mathbf{b}' = \mathbf{e}_{1,n} \alpha + (\mathbf{C}_\ell(p) \mathbf{S}' - \mathbf{S}' \mathbf{C}_r(q)) \mathbf{e}_{1,k}, \]
(6.41)
\[ \mathbf{d}' = \alpha \mathbf{e}_{1,k}^\top - \mathbf{e}_{1,n} (\mathbf{C}_\ell(p) \mathbf{S}' - \mathbf{S}' \mathbf{C}_r(q)) (\mathbf{I}_k - \mathbf{e}_{1,k} \mathbf{e}_{1,k}^\top), \quad \alpha \in \mathbb{F}, \]
and that conversely, if (6.34) holds for some \( \mathbf{b}' \) and \( \mathbf{d}' \), the latter two are necessarily of the form (6.41) for some \( \alpha \in \mathbb{F} \) (the formulas (6.41) can be
made more symmetric upon shifting the parameter α but we do not need this). Now we use the first formula in (6.35) with \( b' \) as in (6.41) to get
\[
f(z) = \mathcal{A}_n(z) e_{1,n} \alpha + \mathcal{A}_n(z)(C_p S' - S' C_q) e_{1,k} + p(z) e_{n,n} S' \mathcal{A}_k(z)^\top.
\]
Replacing in the latter formula \( C_p \), \( C_q \), \( S' \) by \( A, B, S \) according to (6.30), (6.33) and taking into account that \( \mathcal{A}_n(z) e_{1,n} = 1 \), we get (6.39). \( \square \)

Now we drop the two-sided condition (6.29) getting back to the problem TSP, namely: given \( \Omega \) as in (6.28) (without \( S \) though), find an \( f \in \mathbb{F}[z] \) subject to interpolation conditions (6.1).

**Theorem 6.11.** The problem TSP has a solution if and only if the Sylvester equation
\[
AY - YB = bu - vd
\]
adopts a solution \( Y \in \mathbb{F}^{n \times k} \). For each such solution \( Y \), the polynomial
\[
f_Y(z) = \mathcal{A}_n(z) C_{A,v}^{-1} b + \mathcal{P}_{A,v}(z) e_{n,n} C_{A,v}^{-1} Y \mathcal{D}^{-1}_{u,B} \mathcal{A}_k(z)^\top
\]
\[
= d \mathcal{D}^{-1}_{u,B} \mathcal{A}_k(z)^\top + \mathcal{A}_n(z) C_{A,v}^{-1} Y \mathcal{D}^{-1}_{u,B} e_{k,k} \mathcal{P}_{u,B}(z)
\]
(6.43) satisfies conditions (6.1). Moreover either of the formulas (6.43) establishes a one-to-one correspondence between solutions \( Y \) to the equation (6.42) and solutions to the problem TSP.

**Proof.** As in the previous proof, we pass to the equivalent interpolation problem with interpolation conditions (6.31), (6.32). We next multiply both sides of (6.42) by \( C_{A,v}^{-1} \) on the left and by \( \mathcal{D}^{-1}_{u,B} \) on the right. On account of (6.30) and the rightmost definitions in (6.31), (6.32), the resulting equality can be written as
\[
C_p X - XC_q = b' e_{1,k}^\top - e_{1,n} d', \quad \text{where} \quad X = C_{A,v}^{-1} Y \mathcal{D}^{-1}_{u,B}.
\]
(6.44)
Since \( Y \) solves the Sylvester equation (6.42) if and only if \( X = C_{A,v}^{-1} Y \mathcal{D}^{-1}_{u,B} \) solves (6.44), all the statements now follow from Theorem 6.6. The formulas for \( f_X \) are the same as in (6.35) (but with \( X \) instead of \( S' \)). Writing these formulas in terms of \( \Omega \) and \( Y \) (rather than \( \Omega' \) and \( X \)), again making use of (6.30) and the rightmost equalities in (6.31), (6.32), we get (6.43). \( \square \)

**Remark 6.12.** Note that the actual parameters in formulas (6.43) are the bottom row and the rightmost column of the matrix \( C_{A,v}^{-1} Y \mathcal{D}^{-1}_{u,B} \) rather than the whole matrix \( Y \). In other words, the number of independent scalar parameters in the parametrization formulas (6.43) is at most \( \min\{n, k\} \).

**Remark 6.13.** The polynomial \( f_Y \) defined in (6.43) also can be written as
\[
f_Y(z) = \mathcal{A}_n(z) C_{A,v}^{-1} \left( b + (z I_n - A) Y \mathcal{D}^{-1}_{u,B} \mathcal{A}_k(z)^\top \right)
\]
\[
= \left( d + \mathcal{A}_n(z) C_{A,v}^{-1} (z I_k - B) \right) \mathcal{D}^{-1}_{u,B} \mathcal{P}_{u,B}(z).
\]
(6.45)
Proposition 6.14. Given the data set $TSP$ algebraic and let $F$ be polynomial. If the matrix

$$A\mathcal{C}_{A,v} - \mathcal{C}_{A,v}F_n = A^n\mathcal{v}e_{n,n}^\top, \quad \mathcal{D}_{u,B}B - F_k^\top \mathcal{D}_{u,B} = e_{k,k}uB^k.$$  
(6.46)

where $F_n = J_n(0)$ is the $n \times n$ lower triangular Jordan block with zeros on the main diagonal. Indeed, by the definition (3.1) of $\mathcal{C}_{A,v}$, we have

$$A\mathcal{C}_{A,v} - \mathcal{C}_{A,v}F_n = [Av \ldots A^{n-1}v \quad A^n\mathcal{v}] - [Av \ldots A^{n-1}v \quad 0]$$
$$= [0 \ldots 0 \quad A^n\mathcal{v}] = A^n\mathcal{v}e_{n,n}^\top,$$

verifying the first equality in (6.46). The second follows similarly from the definition (3.2). Making use of equalities (6.46) along with the identity

$$z^n e_{n,n}^\top + \mathfrak{A}_n(z)F_n = z\mathfrak{A}_n(z)$$
(see (1.1)) and explicit formulas (3.3) and (3.4) of $\mathcal{P}_{A,v}$ and $\mathcal{P}_{u,B}$, we get

$$\mathcal{P}_{A,v}(z)e_{n,n}^\top \mathcal{C}_{A,v}^{-1} = z^n e_{n,n}^\top \mathcal{C}_{A,v}^{-1} - \mathfrak{A}_n(z)A^n\mathcal{v}e_{n,n}^\top \mathcal{C}_{A,v}^{-1}$$
$$= z^n e_{n,n}^\top \mathcal{C}_{A,v}^{-1} - \mathfrak{A}_n(z)(A\mathcal{C}_{A,v} - \mathcal{C}_{A,v}F_n)\mathcal{C}_{A,v}^{-1}$$
$$= (z^n e_{n,n}^\top - \mathfrak{A}_n(z)F_n)\mathcal{C}_{A,v}^{-1} - \mathfrak{A}_n(z)\mathcal{C}_{A,v}^{-1}A$$
$$= z\mathfrak{A}_n(z)\mathcal{C}_{A,v}^{-1} - \mathfrak{A}_n(z)\mathcal{C}_{A,v}^{-1}A$$
$$= \mathcal{A}_n(z)\mathcal{C}_{A,v}^{-1}(zI_n - A).$$

and similarly,

$$\mathcal{D}_{u,B}^{-1}e_{k,k}\mathcal{P}_{u,B}(z) = (zI_k - B)\mathcal{D}_{u,B}^{-1}\mathfrak{A}_k(z)^\top.$$

Substituting the two latter equalities into (6.43), we arrive at (6.45). \qed

As an application of Theorem 6.11, we get simple sufficient conditions for the problem $TSP$ to have a unique solution.

**Proposition 6.14.** Given the data set (6.28), let us assume that $A$ is algebraic and let $\mu_A(z) = z^K + \mu_{K-1}z^{K-1} + \ldots + \mu_0$ be its minimal central polynomial. If the matrix $\mu_A(B)$ is invertible, then the problem $TSP$ has a unique solution given by formulas (6.43) with

$$Y = \sum_{i=1}^\kappa \mu_j \sum_{i=0}^{j-1} A^i(\mathcal{v}d - \mathcal{b}u)B^{j-i-1} \cdot \mu_A(B)^{-1}.$$  
(6.47)

**Proof.** Since $\mu_A \in \mathbb{F}[z]$ and $\mu_A(A) = 0$, we have for any $Y \in \mathbb{F}^{n \times k}$,

$$-Y\mu_A(B) = \mu_A(A)Y - Y\mu_A(B) = \sum_{j=1}^\kappa \mu_j \sum_{i=0}^{j-1} A^i(AY - YB)B^{j-i-1}.$$  
(6.48)

If $Y$ satisfies (6.42), we replace $AY - YB$ on the right side of (6.48) by $\mathcal{b}u - \mathcal{v}d$ and see that $Y$ is uniquely defined from (6.48) by the formula
(6.47). To verify that $Y$ of the form (6.47) indeed satisfies (6.42), we use equality (6.48) with $v \mathbf{d} - b \mathbf{u}$ instead of $Y$:

$$(b \mathbf{u} - v \mathbf{d}) \mu_A(B) = \sum_{j=1}^{\kappa} \mu_j \sum_{i=0}^{j-1} A^i(Av - b \mathbf{u}) - (v \mathbf{d} - b \mathbf{u})B^{j-i-1}.$$ 

If $Y$ is defined as in (6.47), the expression on the right side can be written as $AY \mu_A(B) - Y \mu_A(B)$. Since the matrices $B$ and $\mu_A(B)$ commute, we therefore have

$$(b \mathbf{u} - v \mathbf{d}) \mu_A(B) = AY \mu_A(B) - Y \mu_A(B),$$

which is equivalent to (6.42), since $\mu_A(B)$ is invertible. The rest follows by Theorem 6.11. □

6.5. Lagrange interpolation. Given interpolation nodes $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_k$ in $F$ along with target values $b_1, \ldots, b_n, d_1, \ldots, d_k$, the two-sided Lagrange interpolation problem consists of finding an $f \in F[z]$ such that

$$f^{\ell}(\alpha_i) = b_i \quad (i = 1, \ldots, n) \quad \text{and} \quad f^{\mathbf{r}}(\beta_j) = d_j \quad (j = 1, \ldots, k). \quad (6.49)$$

We refer to [2] for a detailed treatment of this problem. Here we only show that under the assumption that

(a) the set $\Lambda^{\ell} = \{\alpha_1, \ldots, \alpha_n\}$ is left $P$-independent,
(b) the set $\Lambda^{\mathbf{r}} = \{\beta_1, \ldots, \beta_k\}$ is right $P$-independent, \quad (6.50)

the problem can be embedded into the scheme of TSP. To this end, note that interpolation conditions (5.2) and (5.3) specified to the case

$$A = \begin{bmatrix} \alpha_1 & 0 \\ \vdots & \vdots \\ 0 & \alpha_n \end{bmatrix}, \quad B = \begin{bmatrix} \beta_1 & 0 \\ \vdots & \vdots \\ 0 & \beta_k \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad (6.51)$$

amount to conditions (6.49). By Proposition 3.9, the assumptions (6.50) ensure the pair $(A, \mathbf{v})$ be controllable and the pair $(\mathbf{u}, B)$ be observable and hence, all general results from Section 4.5 apply. Theorem 6.11 describes all solutions to the problem (6.49) in terms of solutions $Y = [y_{ij}]$ of the Sylvester equation (6.42), which in the present setting breaks up into the system of $nk$ scalar equations

$$\alpha_i y_{ij} - y_{ij} \beta_j = b_i - d_j \quad (1 \leq i \leq n, \ 1 \leq j \leq k). \quad (6.52)$$

In the case where $\alpha_i$ or $\beta_j$ are algebraic over $Z_F$, the solvability criterion for the equation (6.52) (as well as the parametrization of all solutions in the indeterminate case) known from [9] lead to an explicit description of all solutions to the problem (6.49).

Acknowledgements: The project was partially supported by by Simons Foundation grant 524539
Declaration of competing interest: The author declared that he had no conflicts of interest with respect to their authorship or the publication of this article.

References

1. G. Birkhoff, *The algebra of multivariate interpolation*, in: *Constructive approaches to mathematical models*, pp. 345–363, Academic Press, Ont., 1979.
2. V. Bolotnikov, *Lagrange interpolation over division rings*, Comm. Algebra 48 (2020), no. 9, 4065–4084.
3. P. M. Cohn, *Free rings and their relations*, Academic Press, London, 1971.
4. P. M. Cohn, *The similarity reduction of matrices over a skew field*, Math. Z. 132 (1973) 151–163.
5. P. M. Cohn, *Skew fields. Theory of general division rings*, Encyclopedia of Mathematics and its Applications 57, Cambridge University Press, Cambridge, 1995.
6. H. Fitting, *Über den Zusammenhang zwischen dem Begriff der Gleichartigkeit zweier Ideale und dem Äquivalenzbegriff der Elementarteilertheorie*, Math. Ann. 112 (1936), no. 1, 572–582.
7. B. Gordon and T. S. Motzkin, *On the zeros of polynomials over division rings*, Trans. Amer. Math. Soc., 116 (1965) 218–226.
8. N. Jacobson, *The Theory of Rings*, American Mathematical Society, New York, 1943.
9. N. Jacobson, *The equation $x' = xd - dx = b$*, Bull. Amer. Math. Soc. 50, (1944). 902–905.
10. R. E. Kalman, *Contributions to the theory of optimal control*, Bol. Soc. Mat. Mexicana 5 (1960), 102–119.
11. T. Y. Lam, *A general theory of Vandermonde matrices*, Exposition. Math. 4 (1986), no. 3, 193–215.
12. T. Y. Lam and A. Leroy, *Vandermonde and Wronskian matrices over division rings*, J. Algebra 119 (1988), no. 2, 308–336.
13. T. Y. Lam and A. Leroy, *Algebraic conjugacy classes and skew polynomial rings*, in *Perspectives in ring theory*, pp. 153–203, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 233, Kluwer Acad. Publ., Dordrecht, 1988.
14. T. Y. Lam and A. Leroy, *Wedderburn polynomials over division rings. I*, J. Pure Appl. Algebra 186 (2004), no. 1, 43–76.
15. T. Y. Lam, A. Leroy and A. Ozturk, *A. Wedderburn polynomials over division rings. II*, in: *Noncommutative rings, group rings, diagram algebras and their applications*, pp. 73–98, Contemp. Math. 456, Amer. Math. Soc., Providence, RI, 2008.
16. O. Ore, *Theory of non-commutative polynomials*, Ann. of Math. 34 (1933), no. 3, 480–508.
17. L. Solomon, *Similarity of the companion matrix and its transpose. With an appendix by Robert M. Guralnick*, Linear Algebra Appl. 302/303 (1999), 555–561.
18. O. Steinfeld, *Quasi-ideals in rings and semigroups*, Hung. Math. Investigations, 10. Akadémiai Kiadó, Budapest, 1978.
19. N.Wiegmann, *Some theorems on matrices with real quaternion elements*, Canad. J. Math. 7 (1955) 191–201.
20. J. H. M. Wedderburn, *Lectures on Matrices*, Am. Math. Soc., Colloq. Publ. 17. Providence, RI, 1934.

Department of Mathematics, William and Mary, Williamsburg, VA 23187-8795, USA