On the Factorization of Non-Commutative Polynomials (in Free Associative Algebras)

Konrad Schrempf∗

February 26, 2018

Abstract

We describe a simple approach to factorize non-commutative (nc) polynomials, that is, elements in free associative algebras (over a commutative field), into atoms (irreducible elements) based on (a special form of) their minimal linear representations. To be more specific, a correspondence between factorizations of an element and upper right blocks of zeros in the system matrix (of its representation) is established. The problem is then reduced to solving a system of polynomial equations (with at most quadratic terms) with commuting unknowns to compute appropriate transformation matrices (if possible).

Keywords: free associative algebra, factorization of polynomials, minimal linear representation, companion matrix, free field, non-commutative formal power series

AMS Classification: Primary 16K40, 16Z05; Secondary 16G99, 16S10

Introduction

Considering non-commutative (nc) polynomials (elements of the free associative algebra) as elements in its universal field of fractions (free field) seems—at a first glance—like to take a sledgehammer to crack a nut. It is maybe more common to view them as nc rational series (or nc formal power series). Therefore one could skip the rather complicated theory behind free fields, briefly introduced in [Coh03, Section 9.3]. For details we refer to [Coh95, Section 6.4].

However, as we shall see, the normal form of Cohn and Reutenauer [CR94] provides new insights. For a given element (in the free field) minimal linear representations can be transformed into each other by invertible matrices over the ground field.

∗Contact: schrempf@math.tugraz.at, Department of Discrete Mathematics (Noncommutative Structures), Graz University of Technology. Supported by the Austrian FWF Project P25510-N26 “Spectra on Lamplighter groups and Free Probability”
In our case we are interested in the (finite set of) transformation matrices which yield all possible factorizations (modulo insertion of units and permutation of commuting factors) of a given polynomial.

Here we use the concept of admissible linear systems (ALS for short) [Coh72], closely related to linear representations. At any step, an ALS (for a nc polynomial) can be easily transformed into a proper linear system [SS78, Section II.1]. For a very brief introduction within our framework see [Sch17, Section 3].

A more general concept for the factorization of arbitrary elements in free fields (for example non-commutative rational formal power series) is considered in future work. There (left and right) divisors will be defined on the level of linear representations. In this way additional structure compensates what is missing in fields (when every non-zero element is invertible).

Section 1 provides the notation and the basic (algebraic) setup. Section 2 contains the main result, Theorem 2.14, which describes a correspondence between factorizations and upper right blocks of zeros (in the system matrix). Section 3 generalizes the concept of companion matrices to provide immediately minimal linear representations for a broader class of elements in the free associative algebra. Section 4 illustrates the concept step by step including remarks how to (easily) obtain minimal linear representations for polynomials.

To our knowledge the literature on the factorization of elements in free associative algebras is rather sparse. Caruso [Car10] describes ideas of J. Davenport using homogenization. We have not yet investigated how their approach compares to ours. Some special cases (for example variable disjoint factorizations) are treated in [ARJ15].

Here we do not consider factorizations in skew polynomial rings (or rather domains), or —more general— rings satisfying the Ore condition [Coh03, Section 7.1] or [Coh85, Section 0.8]. A starting point in this context is [HL13]. Factorizations of skew polynomials have various connections to other areas [Ret10] and [GRW01], just to mention two.

1 Representing Non-Commutative Polynomials

There are much simpler ways for representing elements in free associative algebras (in such a way that addition and multiplication can be defined properly) than the following general presented here. However, including the inverse we keep the full flexibility, which could be used later to compute the left (or right) greatest common divisor of two polynomials $p, q$ by minimizing the linear representation for $p^{-1}q$ in which common left factors would cancel.

Notation. The set of the natural numbers is denoted by $\mathbb{N} = \{1, 2, \ldots\}$, that including zero by $\mathbb{N}_0$. Zero entries in matrices are usually replaced by (lower) dots to stress the structure of the non-zero entries unless they result from transformations where there were possibly non-zero entries before. We denote by $I_n$ the identity matrix of size $n$ respectively $I$ if the size is clear from the context.
Let \( \mathbb{K} \) be a \textit{commutative} field, \( \overline{\mathbb{K}} \) its algebraic closure and \( X = \{x_1, x_2, \ldots, x_d\} \) be a \textit{finite} alphabet. \( \mathbb{K}(X) \) denotes the free associative algebra or free \( \mathbb{K} \)-algebra (or “algebra of non-commutative polynomials”) and \( \mathbb{K}(\langle X \rangle) \) denotes the \textit{universal field of fractions} (or “free field”) of \( \mathbb{K}(X) \) [Coh95], [CR99]. In our examples the alphabet is usually \( X = \{x, y, z\} \). Including the algebra of nc rational series \( \mathbb{K}^{\text{rat}} \langle \langle X \rangle \rangle \) we have the following chain of inclusions: \( \mathbb{K} \subseteq \mathbb{K}(X) \subseteq \mathbb{K}^{\text{rat}} \langle \langle X \rangle \rangle \subseteq \mathbb{K}(\langle X \rangle) \).

The \textit{free monoid} \( X^* \) generated by \( X \) is the set of all \textit{finite words} \( x_{i_1}x_{i_2} \cdots x_{i_m} \) with \( i_k \in \{1, 2, \ldots, d\} \). An element of the alphabet is called \textit{letter}, an element of \( X^* \) is called \textit{word}. The multiplication on \( X^* \) is just the \textit{concatenation} of words, that is, \( (x_{i_1} \cdots x_{i_m}) \cdot (x_{j_1} \cdots x_{j_n}) = x_{i_1} \cdots x_{i_m}x_{j_1} \cdots x_{j_n} \), with neutral element \( 1 \), the \textit{empty word}. The \textit{length} of a word \( w = x_{i_1}x_{i_2} \cdots x_{i_m} \) is \( m \), denoted by \( |w| = m \) or \( \ell(w) = m \).

For detailed introductions see [BR11, Chapter 1] or [SS78, Section I.1].

**Definition 1.1** (Inner Rank, Full Matrix [Coh85], [CR99]). Given a matrix \( A \in \mathbb{K}(X)^{n \times n} \), the \textit{inner rank} of \( A \) is the smallest possible dimension among all linear representations of \( f \). The \textit{dimension} of the representation is \( \dim f = n \). It is called \textit{minimal} if \( A \) has the smallest possible dimension among all linear representations of \( f \). A minimal one \( \pi_f \) defines the \textit{rank} of \( f \), \( \text{rank} f = \dim f \). The “empty” representation \( \pi = (,,) \) is the minimal representation for \( f = 0 \) in \( \mathbb{K}(\langle X \rangle) \) with \( \dim = 0 \).

**Definition 1.2** (Linear Representations, Dimension, Rank [CR94], [CR99]). Let \( f \in \mathbb{K}(\langle X \rangle) \). A \textit{linear representation} of \( f \) is a triple \( \pi_f = (u, A, v) \) with \( u \in \mathbb{K}^{1 \times n} \), full \( A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d, A_\ell \in \mathbb{K}^{n \times n} \) and \( v \in \mathbb{K}^n \) such that \( f = uA^{-1}v \). The \textit{dimension} of the representation is \( \dim \pi_f = n \). It is called \textit{minimal} if \( A \) has the smallest possible dimension among all linear representations of \( f \). A minimal one \( \pi_f \) defines the \textit{rank} of \( f \), \( \text{rank} f = \dim \pi_f \). The “empty” representation \( \pi = (,,) \) is the minimal representation for \( f = 0 \) in \( \mathbb{K}(\langle X \rangle) \) with \( \dim = 0 \).

**Definition 1.3** (Left and Right Families [CR94]). Let \( \pi = (u, A, v) \) be a linear representation of \( f \in \mathbb{K}(\langle X \rangle) \) of dimension \( n \). The families \( \{s_1, s_2, \ldots, s_n\} \subseteq \mathbb{K}(\langle X \rangle) \) with \( s_i = (A^{-1}v)_i \) and \( \{t_1, t_2, \ldots, t_n\} \subseteq \mathbb{K}(\langle X \rangle) \) with \( t_j = (uA^{-1})_j \) are called \textit{left family} and \textit{right family} respectively. \( L(\pi) = \text{span}\{s_1, \ldots, s_n\} \) and \( R(\pi) = \text{span}\{t_1, \ldots, t_n\} \) denote their linear spans.

**Proposition 1.4** ([CR94], Proposition 4.7). A representation \( \pi = (u, A, v) \) of an element \( f \in \mathbb{K}(\langle X \rangle) \) is \textit{minimal} if and only if both, the left family and the right family are \( \mathbb{K} \)-linearly independent.

**Definition 1.5** (Admissible Linear Systems [Coh72], Admissible Transformations [Sch17]). A linear representation \( \mathcal{A} = (u, A, v) \) of \( f \in \mathbb{K}(\langle X \rangle) \) is called \textit{admissible linear system} (for \( f \)), denoted by \( \mathcal{A}s = v \), if \( u = 1 \). The element \( f \) is then the first component of the (unique) solution vector \( s \). Given a linear representation \( \mathcal{A} = (u, A, v) \) of dimension \( n \) of \( f \in \mathbb{K}(\langle X \rangle) \) and invertible matrices \( P, Q \in \mathbb{K}^{n \times n} \), the transformed \( P\mathcal{A}Q = (uQ, P\mathcal{A}Q, Pv) \) is again a linear representation (of \( f \)). If \( \mathcal{A} \) is an ALS, the transformation \( (P, Q) \) is called \textit{admissible} if the first row of \( Q \) is \( \mathcal{A}s = v \). Transformations can be done by elementary row- and column operations, explained in detail in [Sch17, Remark 1.12]. For further remarks and connections to the related concepts of linearization and realization see [Sch17, Section 1].

3
For rational operations on ALS level see the following proposition. If an ALS is minimal then more refined versions of an inverse give a minimal ALS again. For a detailed discussion we refer to [Sch17, Section 4].

**Proposition 1.6** (Rational Operations [Sch17, Proposition 1.13]). Let \(0 \neq f, g \in \mathbb{K}((X))\) be given by the admissible linear systems \(A_f = (u_f, A_f, v_f)\) and \(A_g = (u_g, A_g, v_g)\) respectively and let \(0 \neq \mu \in \mathbb{K}\). Then admissible linear systems for the rational operations can be obtained as follows:

The scalar multiplication \(\mu f\) is given by
\[
\mu A_f = (u_f, A_f, \mu v_f).
\]
The sum \(f + g\) is given by
\[
A_f + A_g = \left(\begin{array}{c} u_f \\ A_f \\ v_f \end{array}\right) + \left(\begin{array}{c} A_f \mu A_f \\ A_g \\ v_g \end{array}\right).
\]
The product \(fg\) is given by
\[
A_f \cdot A_g = \left(\begin{array}{c} u_f \\ A_f \\ v_f \end{array}\right) \cdot \left(\begin{array}{c} A_f \\ A_g \\ v_g \end{array}\right).
\]
And the inverse \(f^{-1}\) is given by
\[
A_f^{-1} = \left(\begin{array}{c} 1 \\ -v_f \\ A_f \\ u_f \end{array}\right).
\]

**Definition 1.7.** An element in \(\mathbb{K}((X))\) is called regular, if it has a linear representation \((u, A, v)\) with \(A = I - M\), where \(M = M_1 \otimes x_1 + \ldots + M_d \otimes x_d\) with \(M_i \in \mathbb{K}^{n \times n}\) for some \(n \in \mathbb{N}\), that is, \(A_0 = I\) in Definition 1.2, or equivalently, if \(A_0\) is regular (invertible).

**Remark.** \(A = I - M\) is also called a monic pencil [HV07]. A regular element can also be represented by a proper linear system \(s = v + Ms\) [SS78, Section II.1]. For a polynomial \(p \in \mathbb{K}([X]) \subseteq \mathbb{K}^{\text{rat}}([X]) \subseteq \mathbb{K}((X))\) the rank of \(p\) is just the Hankel rank, that is, the rank of its Hankel matrix. See [Fli74] and [SS78, Section II.3].

**Example.** The Hankel matrix for \(p = x(1 - yx) = x - xyx\), with row indices \([1, x, xy, xyx]\) and column indices \([1, x, yx, xyx]\), is
\[
\mathcal{H}(p) = \begin{bmatrix}
1 & 1 & -1 \\
-1 & 1 & 1 \\
-1 & -1 & 1 \\
1 & 1 & 1
\end{bmatrix}.
\]
Its rank is 4. Thus \(\text{rank } p = 4\) and therefore the dimension of any minimal admissible linear system is 4, as will be seen later in the ALS \((2.8)\) in Example 2.7.
The following definitions follow mainly [BS15] and are streamlined to our purpose. We do not need the full generality here. While there is a rather uniform factorization theory in the commutative setting [GHK06, Section 1.1], even the “simplest” non-commutative case, that is, a “unique” factorization domain like the free associative algebra, is not straightforward. For a general (algebraic) point of view we recommend the survey [Sme15]. Factorization in free ideal rings (FIRs) is discussed in detail in [Coh85, Chapter 3]. FIRs play an important role in the construction of free fields. More on “non-commutative” factorization can be found in [Jor89] and [BHL17] (just to mention a few) and the literature therein.

**Definition 1.8** (Similar Right Ideals, Similar Elements [Coh85, Section 3.2]). Let \( R \) be a ring. Two right ideals \( a, b \subseteq R \) are called similar, written as \( a \sim b \), if \( R/a \cong R/b \) as right \( R \)-modules. Two elements \( p, q \in R \) are called similar if their right ideals \( pR \) and \( qR \) are similar, that is, \( pR \sim qR \). See also [Sme15, Section 4.1].

**Definition 1.9** (Left and Right Coprime Elements [BS15, Section 2]). Let \( R \) be a domain and \( H = R^* = R \setminus \{0\} \). An element \( p \in H \) left divides \( q \in H \), written as \( p \mid q \), if \( q \in pH = \{ ph \mid h \in H \} \). Two elements \( p, q \) are called left coprime if for all \( h \) such that \( h \mid p \) and \( h \mid q \) implies \( h \in H^\times = \{ f \in H \mid f \text{ is invertible} \} \), that is, \( h \) is an element of the group of units. Right division \( p \righthalfarrow q \) and the notion of right coprime is defined in a similar way. Two elements are called coprime if they are left and right coprime.

**Definition 1.10** (Atomic Domains, Irreducible Elements [BS15, Section 2]). Let \( R \) be a domain and \( H = R^* \). An element \( p \in H \setminus H^\times \), that is, a non-zero non-unit (in \( R \)), is called an atom (or irreducible) if \( p = q_1q_2 \) with \( q_1, q_2 \in H \) implies that either \( q_1 \in H^\times \) or \( q_2 \in H^\times \). The set of atoms in \( R \) is denoted by \( A(R) \). The (cancellative) monoid \( H \) is called atomic if every non-unit can be written as a finite product of atoms of \( H \). The domain \( R \) is called atomic if the monoid \( R^* \) is atomic.

**Remark.** Similarity of two elements \( a, a' \) in a weak Bezout ring \( R \) is equivalent to the existence of \( b, b' \in R \) such that \( ab' = ba' \) with \( ab' \) and \( ba' \) coprime. The free associative algebra \( \mathbb{K}\langle X \rangle \) is a weak Bezout ring [Coh63, Proposition 5.3 and Theorem 6.2].

**Example.** The polynomials \( p = (1 - xy) \) and \( q = (1 - yx) \) are similar, because \( px = (1 - xy)x = x - xyx = x(1 - yx) = xq \). See also Example 2.7.

**Example.** In the free monoid \( X^* \) the atoms are just the letters \( x_i \in X \).

**Definition 1.11** (Similarity Unique Factorization Domains [Sme15, Definition 4.1]). A domain \( R \) is called similarity factorial (or a similarity-UFD) if \( R \) is atomic and it satisfies the property that if \( p_1p_2 \cdots p_m = q_1q_2 \cdots q_n \) for atoms (irreducible elements) \( p_i, q_j \in R \), then \( m = n \) and there exists a permutation \( \sigma \in S_m \) such that \( p_i \) is similar to \( q_{\sigma(i)} \) for all \( i \in 1, 2, \ldots, m \).

**Proposition 1.12** ([Coh63, Theorem 6.3]). The free associative algebra \( \mathbb{K}\langle X \rangle \) is a similarity (unique) factorization domain.
2 Factorizing non-commutative Polynomials

Our concept for the factorization of nc polynomials relies on minimal linear representations. How to construct the latter is not discussed here. For an overview see [Sch17, Section 3] or the illustration in Section 4.

Using proper linear systems [SS78, Section II.1] would be slightly too restrictive because the only possible admissible transformations are conjugations (with respect to the system matrix). On the other hand, admissible linear systems are too general. Therefore we define a form that is suitable for our purpose.

**Definition 2.1** (Pre-Standard ALS, Pre-Standard Admissible Transformation). An admissible linear system \( A = (u, A, v) \) of dimension \( n \) with system matrix \( A = (a_{ij}) \) for a non-zero polynomial \( 0 \neq p \in K\langle X \rangle \) is called pre-standard, if

1. \( v = [0, \ldots, 0, \lambda]^\top \) for some \( \lambda \in K \) and
2. \( a_{ii} = 1 \) for \( i = 1, 2, \ldots, n \) and \( a_{ij} = 0 \) for \( i > j \), that is, \( A \) is upper triangular.

A pre-standard ALS is also written as \( A = (1, A, \lambda) \) with \( 1, \lambda \in K \). An admissible transformation \( (P, Q) \) for a pre-standard ALS \( A \) is called pre-standard, if the transformed system \( PAQ \) is again pre-standard.

To be able to proof that the construction in Proposition 2.6 leads to a minimal linear representation (for the product of two nc polynomials) some preparation is necessary. One of the main tools is Lemma 2.2 [Coh95, Corollary 6.3.6]. Although we are working with regular elements only, invertibility of the constant coefficient matrix \( A_0 \) (in the system matrix) does not have to be assumed in Lemma 2.3.

Minimality (of an admissible linear system) is a crucial assumption for the main result, Theorem 2.14 (polynomial factorization).

**Lemma 2.2** ([Coh95, Corollary 6.3.6]). A linear square matrix over \( K\langle X \rangle \) which is not full is associated over \( K \) to a linear hollow matrix.

**Lemma 2.3.** Let \( A = (u, A, v) \) be an ALS of dimension \( n \geq 2 \) with \( K \)-linearly independent left family \( s = A^{-1}v \) and \( B = B_0 \otimes 1 + B_1 \otimes x_1 + \ldots + B_d \otimes x_d \) with \( B_\ell \in K^{m \times n} \), such that \( Bs = 0 \). Then there exists a (unique) \( T \in K^{m \times n} \) such that \( B = TA \).

**Proof.** Without loss of generality assume that \( v = [0, \ldots, 0, 1]^\top \) and \( m = 1 \). Since \( A \) is full and thus invertible over the free field, there exists a unique \( T \) such that \( B = TA \), namely \( T = BA^{-1} \) in \( K\langle X \rangle^{1 \times n} \). The last column in \( T \) is zero because \( 0 = Bs = BA^{-1}v = Tv \). Now let \( A' \) denote the matrix \( A \) whose last row is removed and \( A_B' \) the matrix obtained from \( A \) when the last row is replaced by \( B \). \( A_B' \) cannot be full since \( s \in \ker A_B' \) would give a contradiction: \( s = (A_B')^{-1}0 = 0 \).

The main idea of this proof is the same as in that of [Sch17, Lemma 4.18], that is, to show that there are no invertible matrices \( P' \in K^{(n-1) \times (n-1)} \) and \( Q \in K^{n \times n} \) with \( (Q^{-1}s)_1 = s_1 \), such that \( P'A'Q \) contains a zero block of size \( (n - i) \times i \) for
$i = 1, \ldots, n - 1$. Such a transformation would contradict minimality of $\mathcal{A}$. We omit the details here. Since no zero block of size $(n - i) \times i$ can be constructed, no zero block of size $(n - i + 1) \times i$ is possible, yielding, by Lemma 2.2, that $A'_B$ is associated over $\mathbb{K}$ to a linear hollow matrix with a $1 \times n$ block of zeros, say in the last row (columns are left untouched):

$$
\begin{bmatrix}
I_{n-1} & A' \\
T' & B
\end{bmatrix} I_n = \begin{bmatrix} A' \end{bmatrix}.
$$

The matrix $T = [-T', 0] \in \mathbb{K}^{1 \times n}$ satisfies $B = TA$.

Remark. Although the ALS in Lemma 2.3 does not have to be minimal, $\mathbb{K}$-linear independence of the left family is an important assumption for two reasons. One is connected to “pathological” situations, compare with [Sch17, Example 2.5]. An entry corresponding to some $s_j = 0$, say for $j = 3$, could be arbitrary:

$$
\begin{bmatrix} 1 & -x \\
. & . \\
. & 1
\end{bmatrix} s = \begin{bmatrix} . \\
. \\
1
\end{bmatrix}.
$$

For $B = [2, -2x, y]$ the transformation $T$ has non-scalar entries: $T = [2, yz^{-1}, 0]$. The other reason concerns the exclusion of other possibilities for non-fullness except the last row. For $B = [0, 0, 1]$ the matrix

$$
A'_B = \begin{bmatrix} 1 & -x \\
. & . \\
. & 1
\end{bmatrix}
$$

is hollow. However, the transformation we are looking for here is $T = [0, z^{-1}, 0]$.

Lemma 2.4. Let $\mathcal{A} = (u, A, v)$ be a pre-standard ALS of dimension $n \geq 2$ with $\mathbb{K}$-linearly dependent left family $s = A^{-1}v$. Let $A = (a_{ij})$. Let $m \in \{2, 3, \ldots, n\}$ be the minimal index such that the left subfamily $s = (A^{-1}v)^n_{i=m}$ is $\mathbb{K}$-linearly independent. Then there exists matrices $T, U \in \mathbb{K}^{1 \times (n+1-m)}$ such that

$$
U + (a_{m-1,j})_{j=m}^n - T(a_{ij})_{i,j=m}^n = [0 \ldots 0] \quad \text{and} \quad T(v_i)_{i=m}^n = 0.
$$

Proof. By assumption, the left subfamily $(s_{m-1}, s_m, \ldots, s_n)$ is $\mathbb{K}$-linearly dependent. Thus there are $\kappa_j \in \mathbb{K}$ such that $s_{m-1} = \kappa_ms_m + \kappa_{m+1}s_{m+1} + \ldots + \kappa_ns_n$. Let $U = [\kappa_m, \kappa_{m+1}, \ldots, \kappa_n]$. Then $s_{m-1} - Us = 0$. Since $\mathcal{A}$ is pre-standard, $v_{m-1} = 0$. Now we can apply Lemma 2.3 with $B = U + [a_{m-1,m}, a_{m-1,m+1}, \ldots, a_{m-1,n}]$ (and $s$).

Hence, there exists a matrix $T \in \mathbb{K}^{1 \times (n+1-m)}$ such that

$$
U + [a_{m-1,m} \ldots a_{m-1,n}] - T \begin{bmatrix} 1 & a_{m+1} & \cdots & a_{m,n} \\
. & \cdots & . & . \\
. & 1 & \cdots & . \\
. & . & . & 1
\end{bmatrix} = [0 \ldots 0]. \quad (2.5)
$$

Recall that the last column of $T$ is zero, whence $T(v_i)_{i=m}^n = 0$. 

7
Proposition 2.6 (Minimal Polynomial Multiplication). Let \( 0 \neq p, q \in \mathbb{K}(X) \) be given by the minimal pre-standard admissible linear systems \( A_p = (u_p, A_p, v_p) = (1, A_p, \lambda_p) \) and \( A_q = (u_q, A_q, v_q) = (1, A_q, \lambda_q) \) of dimension \( n_p, n_q \geq 2 \) respectively. Then a minimal ALS for \( pq \) has dimension \( n = n_p + n_q - 1 \) and can be constructed in the following way:

1. Construct the following ALS \( A' = (u', A', v') \) for the product \( pq \):
   \[
   \begin{bmatrix}
   A_p & -v_p u_q \\
   . & A_q
   \end{bmatrix} s' = \begin{bmatrix}
   v_q
   \end{bmatrix}.
   \]

2. Add \( \lambda_p \)-times column \( n_p \) to column \( n_p + 1 \) (the entry \( s'_{n_p} \) becomes zero).

3. Remove row \( n_p \) of \( A' \) and \( v' \) and column \( n_p \) of \( A' \) and \( u' \) and denote the new (pre-standard) ALS of dimension \( n_p + n_q - 1 \) by \( A = (u, A, v) = (1, A, \lambda) \).

Proof. The left family of \( A' \) is
   \[
   s' = \begin{bmatrix}
   A_p^{-1} & A_q^{-1} v_p u_q \\
   . & .
   \end{bmatrix} \begin{bmatrix}
   v_q
   \end{bmatrix} = \begin{bmatrix}
   s_p, q
   \end{bmatrix}.
   \]

See also [Sch17, Proposition 1.13]. Clearly, \( A \) is again pre-standard with \( \lambda = \lambda_q \). Both systems \( A_p \) and \( A_q \) are minimal. Therefore their left and right families are \( \mathbb{K} \)-linearly independent. Without loss of generality assume that \( \lambda_p = 1 \). Then the last entry \( t_{n_p} \) of the right family of \( A_p \) is equal to \( p \). Let \( k = n_p \). We have to show that both, the left and the right family
   \[
   s = (s_1, s_2, \ldots, s_n) = (s_1^p q_1, \ldots, s_{k-1}^p q_1 q, s_2^q, \ldots, s_{n_q}^q),
   \]
   \[
   t = (t_1, t_2, \ldots, t_n) = (t_1^p, \ldots, t_{k-1}^p p t_{q_2}, \ldots, t_{q_{n_q}}^p)
   \]

of \( A \) are \( \mathbb{K} \)-linearly independent. Assume the contrary for \( s \), say there is an index \( 1 < m \leq k \) such that \( (s_{m-1}, s_m, \ldots, s_n) \) is \( \mathbb{K} \)-linear dependent while \( (s_m, \ldots, s_n) \) is \( \mathbb{K} \)-linear independent. Then, by Lemma 2.4 there exist matrices \( T, U \in \mathbb{K}^{1 \times (n-m+1)} \) such that (2.5) holds and therefore invertible matrices \( P, Q \in \mathbb{K}^{n \times n} \),
   \[
   P = \begin{bmatrix}
   I_{m-2} & . & 1 & T \\
   . & . & . & .
   \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix}
   I_{m-2} & . & . \\
   . & 1 & U \\
   . & . & I_{n-m+1}
   \end{bmatrix},
   \]

that yield equation \( s_{m-1} = 0 \) (in row \( m - 1 \)) in \( P A Q \). Let \( \hat{P} \) (respectively \( \hat{Q} \)) be the upper left part of \( P \) (respectively \( Q \)) of size \( k \times k \). Then the equation in row \( m - 1 \) in \( \hat{P} A \hat{Q} \) is \( s_{m-1} = \alpha \in \mathbb{K} \), contradicting \( \mathbb{K} \)-linear independence of the left family of \( A_p \) since \( s_k^p = \lambda_p \in \mathbb{K} \). A similar argument (and a variant of Lemma 2.4) for the right family yields its \( \mathbb{K} \)-linear independence. Hence, by Proposition 1.4, the admissible linear system \( A \) (for \( pq \)) is minimal. \( \square \)
Example 2.7. The polynomials \( p = x \in \mathbb{K}\langle X \rangle \) and \( q = 1 - yx \in \mathbb{K}\langle X \rangle \) have the minimal pre-standard admissible linear systems

\[
A_p = \begin{pmatrix} 1 & 1 -x \\ . & 1 \end{pmatrix}, 1 \) and \( A_q = \begin{pmatrix} 1 & y \\ . & 1 -x \end{pmatrix}, 1 \)
\]
respectively. Then a pre-standard ALS for \( pq = x(1 - yx) \) is given by

\[
\begin{bmatrix} 1 & -x & . & . \\ . & 1 & -1 & . \\ . & . & 1 & y -1 \\ . & . & . & 1 \\
\end{bmatrix} \begin{bmatrix} s_1 \\ . \\ . \\ . \\
\end{bmatrix} = \begin{pmatrix} x(1 - yx) \\ 1 - yx \\ 1 - yx \\ x \\
\end{pmatrix}.
\]

Adding column 2 to column 3 (and subtracting \( s_3 \) from \( s_2 \)) yields

\[
\begin{bmatrix} 1 & -x & -x & . \\ . & 1 & 0 & . \\ . & . & 1 & y -1 \\ . & . & . & 1 \\
\end{bmatrix} \begin{bmatrix} s_1 \\ . \\ . \\ . \\
\end{bmatrix} = \begin{pmatrix} x(1 - yx) \\ 0 \\ 1 - yx \\ x \\
\end{pmatrix}.
\]

thus the minimal pre-standard ALS

\[
A = \begin{pmatrix} 1 -x & 0 & 0 \\ . & 1 & y -1 \\ . & . & 1 -x \\ . & . & . & 1 \\
\end{pmatrix} \begin{bmatrix} . \\ . \\ . \\ . \\
\end{bmatrix}.
\]

(2.8)

Note the upper right \( 1 \times 2 \) block of zeros in the system matrix of \( A \).

Definition 2.9 (Atomic Admissible Linear Systems). A minimal pre-standard ALS \( A = (1, A, \lambda) \) of dimension \( n \geq 2 \) is called atomic (irreducible), if there is no pre-standard admissible transformation \((P, Q)\) such that \( PAQ \) has an upper right block of zeros of size \((n - i - 1) \times i\) for some \( i = 1, 2, \ldots, n - 2 \).

If we take the minimal ALS \((2.8)\) for \( pq = x(1 - yx) \) from Example 2.7 and add column 2 to column 4, we obtain

\[
\begin{bmatrix} 1 & -x & . \\ . & 1 & y \\ . & . & 1 -x \\ . & . & . \\
\end{bmatrix} \begin{bmatrix} s_1 \\ . \\ . \\ . \\
\end{bmatrix} = \begin{pmatrix} x(1 - yx) \\ -yx \\ 1 - yx \\ x \\
\end{pmatrix}.
\]

Subtracting row 3 from row 1 yields

\[
\begin{bmatrix} 1 & -x & 0 \\ . & 1 & y \\ . & . & 1 -x \\ . & . & . \\
\end{bmatrix} \begin{bmatrix} s_1 \\ . \\ . \\ . \\
\end{bmatrix} = \begin{pmatrix} x(1 - yx) \\ -yx \\ 1 - yx \\ x \\
\end{pmatrix}.
\]
with an upper right $2 \times 1$ block of zeros in the system matrix. We would obtain the same system by minimal polynomial multiplication of $1 - xy$ and $x$. This illustrates that we can find the factors of non-commutative polynomials by searching for pre-standard admissible transformations that give a corresponding upper right block of zeros in the transformed system matrix.

Theorem 2.14 will establish the correspondence between a factorization into atoms and the structure of the upper right blocks of zeros. Thus, to factor a polynomial $p$ of rank $n \geq 3$ into non-trivial factors $p = q_1 q_2$ with $\text{rank}(q_i) = n_i \geq 2$ and $n = n_1 + n_2 - 1$, we have to look for (pre-standard admissible) transformations of the form

$$ (P, Q) = \begin{pmatrix} 1 & \alpha_{1,2} & \cdots & \alpha_{1,n-1} & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & \alpha_{n-2,n-1} & \cdots & 1 & 0 \\ 0 & 1 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & \beta_{2,3} & \cdots & \beta_{2,n} \\ \vdots & \ddots & \ddots & \ddots \\ 1 & \cdots & 1 & \vdots \\ \beta_{n-1,n} & \cdots & \cdots & 1 \end{pmatrix} $$

(2.10)

with entries $\alpha_{ij}, \beta_{ij} \in \mathbb{K}$. In general this is a non-linear problem with $(n - 2)(n - 1)$ unknowns.

**Definition 2.11** (Standard Admissible Linear Systems). Every

- **ALS** $(1, I_1, \lambda)$ for a scalar $0 \neq \lambda \in \mathbb{K}$,
- **atomic ALS** and
- **non-atomic minimal** pre-standard ALS $A = (1, A, \lambda)$ for $p \in \mathbb{K}(X)$ of dimension $n \geq 3$ obtained from minimal multiplication (Proposition 2.6) of atomic admissible linear systems for its atomic factors $p = q_1 q_2 \cdots q_m$ (with atoms $q_i$) is called **standard**.

**Example 2.12.** Let $p = x^2 - 2 \in \mathbb{K}(X)$ be given by the minimal ALS

$$ A = \begin{pmatrix} 1 & \cdots & 1 & -x & 2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 1 & -x & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{pmatrix}. $$

If $\mathbb{K} = \mathbb{Q}$ then $A$ is atomic (irreducible). If $\mathbb{K} = \mathbb{R}$ then there is the pre-standard admissible transformation

$$ (P, Q) = \begin{pmatrix} 1 & \sqrt{2} \\ 1 & 1 \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 & -\sqrt{2} \\ \cdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \cdots \\ 1 & \cdots & 1 & 1 \end{pmatrix} $$

such that $A' = PAQ$ is

$$ A' = \begin{pmatrix} 1 & \cdots & 1 & -x + \sqrt{2} & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \cdots \\ 1 & \cdots & 1 & -x - \sqrt{2} & 1 \end{pmatrix}, $$

10
thus $p = x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ in $\mathbb{R}(X)$.

**Remark.** Note that it is easy to check that $p = xy - 2 \in \mathbb{K}(X)$ is atomic, because both entries, $\alpha_{1,2}$ in $P$ and $\beta_{2,3}$ in $Q$ have to be zero (otherwise the upper right entry in $A'$ could not become zero) and hence there is no non-trivial pre-standard admissible transformation, that is, transformation that changes the upper right block structure.

**Lemma 2.13.** Let $0 \neq p, q_1, q_2 \in \mathbb{K}(X)$ be given by the minimal pre-standard admissible linear systems $A = (1, A, \lambda)$, $A_1 = (1, A_1, \lambda_1)$ and $A_2 = (1, A_2, \lambda_2)$ of dimension $n = \text{rank} p$ and $n_1, n_2 \geq 2$ respectively with $p = q_1 q_2$. Then there exists a pre-standard admissible transformation $(P, Q)$ such that $PAQ$ has an upper right block of zeros of size $(n_1 - 1) \times (n_2 - 1) = (n - n_2) \times (n - n_1)$.

**Proof.** Let $A' = (1, A', \lambda_2)$ be the minimal ALS from Proposition 2.6 for $q_1 q_2$. Clearly, we have $\dim A' = n_1 + n_2 - 1 = n = \text{rank} p$. And $A'$ has, by construction, an upper right block of zeros of size $(n_1 - 1) \times (n_2 - 1)$. Both, $A$ and $A'$ represent the same element $p$, thus —by [CR99, Theorem 1.4]— there exists an admissible transformation $(P, Q)$ such that $PAQ = A'$. Since $A$ and $A'$ are pre-standard, $(P, Q)$ is a pre-standard admissible transformation.

**Theorem 2.14 (Polynomial Factorization).** Let $p \in \mathbb{K}(X)$ be given by the minimal pre-standard admissible linear system $A = (1, A, \lambda)$ of dimension $n = \text{rank} p \geq 3$. Then $p$ has a factorization into $p = q_1 q_2$ with $\text{rank}(q_i) = n_i \geq 2$ if and only if there exists a pre-standard admissible transformation $(P, Q)$ such that $PAQ$ has an upper right block of zeros of size $(n_1 - 1) \times (n_2 - 1)$.

**Proof.** If there is such a factorization, then Lemma 2.13 applies. Conversely, if we have such a (pre-standard admissible) transformation, we get the two systems for the factors by “admissibly” inserting a row and a column in the ALS $A$, that is, reversing the steps of Proposition 2.6. See also Example 2.7.

This finishes the algorithm. A simple analogue of [CR99, Theorem 4.1] will do the rest, we do not have to worry about invertibility of the transformation matrices $P$ and $Q$ in (2.10). In our case $\mathbb{K}[\alpha, \beta] = \mathbb{K}[\alpha_{1,2}, \ldots, \alpha_{n-1,n}, \beta_{2,3}, \ldots, \beta_{n-1,n}]$. However, a non-trivial ideal does not guarantee a solution over $\mathbb{K}$. We do not go into details here, how one could test, if there is such a solution (without computing it). For an introduction to Gröbner basis, we recommend [CLO15, Chapter 2].

**Remark.** Note, that in general in order to reverse the multiplication (Proposition 2.6) to find factors, we also need a lower left block of zeros of appropriate size. This important fact is hidden in the pre-standard form of an ALS. A general factorization concept is considered in future work.

**Proposition 2.15.** Let $p \in \mathbb{K}(X)$ be given by the minimal pre-standard admissible linear system $A = (1, A, \lambda)$ of dimension $n = \text{rank} p \geq 3$ and let $(P, Q)$ as in (2.10). Fix a $k \in \{1, 2, \ldots, n-2\}$ and denote by $I_k$ the ideal of $\mathbb{K}[\alpha, \beta]$ generated by the
coefficients of each \( x \in \{1\} \cup X \) in the \((i,j)\) entries of the matrix \( P A Q \) for \( 1 \leq i \leq k \) and \( n - k \leq j \leq n \). Then \( p \) factorizes over \( \mathbb{K}(X) \) into \( p = q_1 q_2 \) with \( \text{rank} q_1 = k + 1 \) and \( \text{rank} q_2 = n - k \) if and only if the ideal \( I_k \) is non-trivial.

Given a polynomial \( p \in \mathbb{K}(X) \) by a minimal pre-standard ALS of dimension \( n = \text{rank} p \geq 2 \) there are at most \( \phi(n) = 2^{n-2} \) (minimal) standard admissible linear systems (with respect to the structure of the upper right blocks of zeros). For \( n = 2 \) this is clear. For \( n > 2 \) the ALS could be atomic or have a block of zeros of size \( 1 \times (n - 2) \) or \( (n - 2) \times 1 \), thus \( \phi(n + 1) = 1 + 2\phi(n) - 1 = 2\phi(n) \) because the system with “finest” upper right structure is counted twice.

Remark. The number of factorizations can be bigger. As an example take the polynomial \( p = (x-1)(x-2)(x-3) \) which has \( 3! = 6 \) different factorizations while the number of standard admissible linear systems is bounded by \( 2^{\text{rank} p - 2} = 4 \).

Let \( p \) be a non-zero polynomial with the factorization \( p = q_1 q_2 \cdots q_m \) into atoms \( q_i \in \mathbb{K}(X) \). Since \( \mathbb{K}(X) \) is a similarity-UFD (Proposition 1.12), each factorization of \( (a \text{ non-zero non-unit}) p \) into atoms has \( m \) factors. Therefore one can define the length of \( p \) by \( \ell(p) = m \). For a word \( w \in X^* \subseteq \mathbb{K}(X) \) the length is \( \ell(w) = |w| \).

By looking at the minimal polynomial multiplication (Proposition 2.6) it is easy to see that the length of an element \( p \in \mathbb{K}(X)^* \) can be estimated by the rank, namely \( \ell(p) \leq \text{rank}(p) - 1 \). More on length functions —and transfer homomorphisms in the context of non-unique factorizations— (in the non-commutative setting) can be found in [Sme15, Section 3] or [BS15].

3 Generalizing the Companion Matrix

For a special case, namely an alphabet with just one letter, the companion matrix of a polynomial \( p(x) \) yields immediately a minimal linear representation of \( p \in \mathbb{K}\langle \{x\} \rangle \). If this is the characteristic polynomial of some (square) matrix \( B \in \mathbb{K}^{m \times m} \), then its eigenvalues can be computed by the techniques from Section 2 (if necessary going over to \( \mathbb{K} \)), illustrated in Example 3.7.

For a broader class of (nc) polynomials, left and right companion systems can be defined. In general minimal pre-standard admissible linear systems are necessary to generalize companion matrices, see Definition 3.6.

**Definition 3.1** (Companion Matrix, Characteristic Polynomial, Normal Form [Gan86, Section 6.6]). Let \( p(x) = a_0 + a_1 x + \ldots + a_{m-1} x^{m-1} + x^m \in \mathbb{K}[x] \). The companion matrix \( L(p) \) is defined as

\[
L(p) = \begin{bmatrix}
0 & 0 & \ldots & 0 & -a_0 \\
1 & 0 & \ddots & \vdots & -a_1 \\
& \ddots & \ddots & 0 & \vdots \\
& & \ddots & 0 & -a_{m-2} \\
& & & 1 & -a_{m-1}
\end{bmatrix}.
\]
Then \( p(x) \) is the characteristic polynomial of \( L = L(p) \):

\[
\det(xI - L) = \det \begin{bmatrix}
x & 0 & \ldots & 0 & a_0 \\
-1 & x & \cdot & \cdot & \cdot & a_1 \\
\cdot & \cdot & \cdot & 0 & \cdot \\
-1 & x & \cdot & \cdot & a_{m-2} \\
-1 & x & \cdot & \cdot & a_{m-1}
\end{bmatrix}.
\]

Given a square matrix \( M \in \mathbb{K}^{m \times m} \), the normal form for \( M \) can be defined in terms of the companion matrix \( L(M) \) of its characteristic polynomial \( p(M) = \det(xI - M) \).

**Remark.** In [Coh95, Section 8.1], \( C(p) = xI - L(p)^\top \) is also called companion matrix. This is justified by viewing \( C(p) \) as linear matrix pencil \( C(p) = C_0 \otimes 1 + C_x \otimes \alpha \). It generalizes nicely for non-commutative polynomials.

Now we leave \( \mathbb{K}[x] = \mathbb{K}\{x\} \) and consider (nc) polynomials \( p \in \mathbb{K}\langle X \rangle \). There are two cases where a minimal ALS can be stated immediately, namely if the support (of the polynomial) can be “built” from the left (in the left family) or from the right (in the right family) with strictly increasing rank. For example, a minimal pre-standard ALS for \( p = a_0 + a_1(x + 2y) + a_2(x - z)(x + 2y) + y(x - z)(x + 2y) \) is given by

\[
\begin{bmatrix}
1 & -y - a_2 & -a_1 & -a_0 \\
-1 & -(x - z) & 1 & -a_0 \\
\cdot & 1 & -(x + 2y) & 1 \\
\cdot & \cdot & \cdot & 1
\end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} p \\ x + 2y \\ 1 \\ 1 \end{bmatrix}.
\]

**Definition 3.2** (Left and Right Companion System). For \( i = 1, 2, \ldots, m \) let \( q_i \in \mathbb{K}\langle X \rangle \) with rank \( q_i = 2 \) and \( a_i \in \mathbb{K} \). For a polynomial \( p \in \mathbb{K}\langle X \rangle \) of the form

\[
p = q_m q_{m-1} \cdots q_1 + a_{m-1} q_{m-1} \cdots q_1 + \ldots + a_2 q_1 + a_1 q_1 + a_0
\]

the pre-standard ALS

\[
\begin{bmatrix}
1 & -q_m & -a_{m-1} & -a_{m-2} & \ldots & -a_1 & -a_0 \\
1 & -q_{m-1} & 0 & \ldots & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
1 & -q_2 & 0 & \cdot & \cdot & \cdot & \cdot \\
1 & 0 & -q_1 & 0 & \cdot & \cdot & \cdot \\
1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot 
\end{bmatrix} s = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}
\]

(3.3)

is called left companion system. And for a polynomial \( p \in \mathbb{K}\langle X \rangle \) of the form

\[
p = a_0 + a_1 q_1 + a_2 q_1 q_2 + \ldots + a_{m-1} q_1 q_2 \cdots q_{m-1} + q_1 q_2 \cdots q_m
\]
the pre-standard ALS
\[
\begin{bmatrix}
1 & -q_1 & 0 & \ldots & 0 & -a_0 \\
1 & -q_2 & \ddots & \vdots & \vdots & -a_1 \\
\vdots & \ddots & \ddots & 0 & \vdots & \vdots \\
1 & -q_{m-1} & 0 & \ldots & -a_{m-2} & -a_{m-1} \\
1 & -q_m & \ldots & \vdots & 1 & -a_m \\
\end{bmatrix}
\]
is called right companion system.

**Proposition 3.5.** For \( i = 1, 2, \ldots, m \) let \( q_i \in \mathbb{K}<X> \) with \( \text{rank } q_i = 2 \). Then the polynomials
\[
p_l = q_m q_{m-1} \cdots q_1 + a_{m-1} q_{m-1} \cdots q_1 + \ldots + a_2 q_1 + a_1 + a_0 \quad \text{and} \quad
p_r = a_0 + a_1 q_1 + a_2 q_2 + \ldots + a_{m-1} q_{m-1} + q_{m-1} + q_1 q_2 \cdots q_m
\]
have rank \( m + 1 \).

**Proof.** Both, the left family \((p, q_{m-1} \cdots q_1, q_2 q_1, q_1, 1)\) and the right family \((1, q_m + a_{m-1}, q_{m-1} + a_{m-2}, \ldots, p)\) for (3.3) are \( \mathbb{K} \)-linearly independent. Thus, the left companion system (for \( p_l \)) is minimal of dimension \( m + 1 \). Hence \( \text{rank } p_l = m + 1 \).

By a similar argument for the right companion system we get \( \text{rank } p_r = m + 1 \).

For a general polynomial \( p \in \mathbb{K}<X> \) with \( \text{rank } p = n \geq 2 \) we can take any *minimal* pre-standard ALS \( \mathcal{A} = (1, A, \lambda) \) to obtain an ALS of the form \((1, A', 1)\) by dividing the last row by \( \lambda \) and multiplying the last column by \( \lambda \). Now we can define a (generalized version of the) companion matrix. Those companion matrices can be used as building blocks to get companion matrices for products of polynomials. This is just a different point of view on the “minimal” polynomial multiplication from Proposition 2.6.

**Remark.** Although nothing can be said in general about minimality of a linear representation for commutative polynomials (in several variables), Proposition 3.5 can be used for constructing minimal linear representations in the commutative case. Because in this case, the rank is the maximum of the ranks of the monomials. For example \( p = x^2 y + xyz = x y x + x y z = x y (x + z) \).

**Definition 3.6** (Companion Matrices). Let \( p \in \mathbb{K}<X> \) with \( \text{rank } p = n \geq 2 \) be given by a minimal pre-standard admissible linear system \( \mathcal{A} = (1, A, 1) \) and denote by \( C(p) \) the upper right submatrix of size \( (n-1) \times (n-1) \). Then \( C(p) \) is called a (nc) companion matrix of \( p \).

**Example 3.7.** Let
\[
B = \begin{bmatrix}
6 & 1 & 3 \\
-7 & 3 & 14 \\
1 & 0 & 1
\end{bmatrix}
\]
Then the characteristic polynomial of \( B \) is \( p(x) = \det(xI - B) = x^3 - 10x^2 + 31x - 30 \).
The left companion system of $p$ is
\[
\begin{bmatrix}
1 & -x + 10 & -31 & 30 \\
1 & -x & . & . \\
. & 1 & -x & . \\
. & . & 1 & -x \\
. & . & . & 1
\end{bmatrix}
\begin{bmatrix}
s_1 \\
s_2 \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}
= \begin{bmatrix}
p(x) \\
x \\
. \\
. \\
. \\
1
\end{bmatrix}.
\]
Applying the transformation $(P, Q)$ with
\[
P = \begin{bmatrix}
1 & 1 & \ldots & 1 \\
-3 & 1 & \ldots & 1 \\
. & 1 & \ldots & 1 \\
. & . & \ldots & 1
\end{bmatrix}
\begin{bmatrix}
1 & 5 \\
1 & 1 \\
1 & 1 \\
1 & 1
\end{bmatrix}
= \begin{bmatrix}
1 & -5 & 6 & . \\
1 & 1 & 3 & . \\
. & 1 & 1 & 1 \\
. & . & . & .
\end{bmatrix}
\]
and
\[
Q = \begin{bmatrix}
1 & 5 & . & . \\
1 & 2 & . & . \\
. & 1 & . & . \\
. & . & 1 & .
\end{bmatrix}
\begin{bmatrix}
1 & . & . \\
1 & . & . \\
1 & . & . \\
1 & . & .
\end{bmatrix}
= \begin{bmatrix}
1 & . & . \\
1 & 5 & 9 & . \\
. & 1 & 3 & . \\
. & . & . & .
\end{bmatrix}
\]
we get the ALS
\[
\begin{bmatrix}
1 & 5 - x & . & . \\
. & 1 & 2 - x & . \\
. & . & 1 & 3 - x \\
. & . & . & 1
\end{bmatrix}
\begin{bmatrix}
s_1 \\
s_2 \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}
= \begin{bmatrix}
p(x) \\
x(3 - 2)(x - 3) \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}.
\]
Thus the eigenvalues of $B$ are 2, 3 and 5. Compare with Example 2.12 for the case when the polynomial does not decompose in linear factors.

4 An Example (step by step)

Let $p = x(1 - yx)(3 - yx)$ and $q = (xy - 1)(xy - 3)x$ be given. Taking companion systems (Definition 3.2) for their factors respectively, by Proposition 2.6 we get the minimal ALS (for $p$):
\[
\begin{bmatrix}
1 & -x & . & . & . \\
1 & -y & -1 & . & . \\
. & 1 & x & . & . \\
. & . & 1 & -y & -3 \\
. & . & . & 1 & x \\
. & . & . & . & 1
\end{bmatrix}
\begin{bmatrix}
s_1 \\
s_2 \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}
= \begin{bmatrix}
x(1 - yx)(3 - yx) \\
(1 - yx)(3 - yx) \\
. \\
. \\
. \\
. \\
3 - yx \\
-3 - yx
\end{bmatrix}
\]
\[
\begin{bmatrix}
x(1 - yx)(3 - yx) \\
(1 - yx)(3 - yx) \\
. \\
. \\
. \\
. \\
3 - yx \\
-3 - yx
\end{bmatrix}
\begin{bmatrix}
s_1 \\
s_2 \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}
= \begin{bmatrix}
x(3 - 2)(x - 3) \\
x(3 - 2)(x - 3) \\
. \\
. \\
. \\
. \\
1
\end{bmatrix}.
\]

Clearly, $p = xyxyx - 4xyx + 3x = q$ and rank $p = rank q = 6$. 

15
Constructing a minimal ALS

If a minimal ALS (for some polynomial) cannot be stated directly by a left or right companion system (Definition 3.2), an ALS for \( p \) will be constructed using rational operations (Proposition 1.6) out of monomials. Then (after suitable transformation of the system matrix into the form \( A = I - M \)) a minimal linear representation can be obtained by the algorithm in [CC80]. Alternatively, it can be constructed directly out of the Hankel matrix by the Extended Ho-Algorithm [FM80]. From a linear representation it is easy to get a pre-standard ALS.

Rather than describing precisely an algorithm which works directly on pre-standard admissible linear systems (keeping the form), we illustrate the steps in detail to get a minimal ALS for the difference of 3\( x \) and 4\( xyx \) given by minimal systems [Sch17, Proposition 4.1]. Optimal runtime complexity is not an issue here, since in general the factorization involves solving a system of non-linear equations which is usually the bottleneck. Additionally—in the spirit of Proposition 2.6—a “polynomial addition” (in general not minimal) could be defined in such a way that \( \dim(A_f + A_g) = n_f + n_g - 2 \) for \( f + g \) with rank \( f, g \geq 2 \), compare with [CR99, Proof of Theorem 2.3].

Recall that an ALS \( A = (u, A, v) \) is minimal if and only if both the left family \( s = A^{-1}v \) is \( \mathbb{K} \)-linearly independent and the right family \( t = uA^{-1} \) is \( \mathbb{K} \)-linearly independent (Proposition 1.4). Let \( f = 3x - 4xyx \). An ALS for \( f \) of dimension \( n = 6 \) is given by

\[
\begin{bmatrix}
1 & -x & \ldots & -1 \\
1 & -y & \ldots & \ldots \\
\ldots & 1 & -x & \ldots \\
\ldots & \ldots & 1 & -x \\
\ldots & \ldots & \ldots & 1
\end{bmatrix}
\begin{bmatrix}
s \\
\end{bmatrix}
\begin{bmatrix}
3x - 4xyx \\
-4yx \\
-4x \\
3x \\
3
\end{bmatrix}.
\]

Note that the free associative algebra \( \mathbb{K}(X) \) is a vector space with basis \( X^* \). Here \( X^* = \{1, x, y, xx, xy, yx, yy, \ldots \} \). Therefore, we can write the left family \( s \) as a matrix (of coordinate row vectors) \( S \in \mathbb{K}^{n \times m_s} \) with column indices \( \{1, x, yx, xyx\} \) and the right family \( t \) as a matrix (of coordinate column vectors) \( T \in \mathbb{K}^{m_r \times n} \) with row indices \( \{1, x, xy, xyx\} \), that is,

\[
S = \begin{bmatrix}
. & 3 & \ldots & -1 \\
. & \ldots & -4 \\
-4 & \ldots & \ldots \\
\end{bmatrix}
\quad\text{and}\quad
T = \begin{bmatrix}
1 & \ldots & 1 \\
1 & \ldots & 1 \\
\ldots & 1 & \ldots \\
\end{bmatrix}.
\]

Both, \( S \) and \( T \) have rank \( 4 < n = 6 \), hence \( A \) cannot be minimal. The entries of the left (respectively right) family are just rows in \( S \), also denoted by \( s_i \) for \( i = 1, \ldots, n \) (respectively columns in \( T \), denoted by \( t_j \) for \( j = 1, \ldots, n \)).
Adding a row $s_k$ to $s_i$ results in subtracting column $i$ from column $k$ in the system matrix $A$. In order to keep the triangular structure $i < k$ must hold. Similarly, adding a column $t_k$ to $t_j$ results in subtracting row $j$ to row $k$, hence $j > k$ must hold. If we want to construct zeros in row $i = 3$ (we cannot produce zeros in row 1 and 2) in $S$ we need to find an (invertible) transformation matrix $Q \in \mathbb{K}^{n \times n}$ of the form (note that we will apply $Q^{-1}$ from the left)

$$
Q = \begin{bmatrix}
1 & \ldots & \\ 1 & \ldots & \beta_4 \beta_5 \beta_6 \\
1 & \ldots & \\
1 & \\
1 & \\
1 & 
\end{bmatrix}
$$

by solving the (underdetermined) linear system with the system matrix consisting of the rows $i + 1, i + 2, \ldots, n$ and the right hand side consisting of the row $i$ of $S$:

$$
[\beta_4 \beta_5 \beta_6] \begin{bmatrix}
-4 \\
3 \\
3 
\end{bmatrix} = \begin{bmatrix}
-4 \\
3 
\end{bmatrix}.
$$

Here $\beta_5 = -4/3$ and we choose $\beta_4 = \beta_6 = 0$. Now row 3 in $Q^{-1}S$ is zero

$$
Q^{-1}S = \begin{bmatrix}
1 & \ldots & . \\
1 & \ldots & . \\
1 & 0 & 4/3 & 0 \\
1 & \ldots & . \\
1 & \ldots & . \\
1 & . & 3 & . \\
1 & 3 & . & . \\
1 & 3 & . & . \\
\end{bmatrix} \begin{bmatrix}
3 & \ldots & -1 \\
. & . & \ldots & -4 \\
. & . & . & 0 \\
. & . & \ldots & -4 \\
. & . & . & . \\
. & . & . & 3 \\
. & . & . & . \\
. & . & . & . \\
\end{bmatrix} = \begin{bmatrix}
3 & \ldots & -1 \\
. & . & \ldots & -4 \\
. & . & . & 0 \\
. & . & \ldots & -4 \\
. & . & . & . \\
. & . & . & 3 \\
. & . & . & . \\
. & . & . & . \\
\end{bmatrix},
$$

that is $s_3' = (Q^{-1}s)_3 = 0$. Therefore we can remove row 3 and column 3 (note that the corresponding entry $u_3$ in $u$ is also zero) from the (admissibly) modified system $(uQ, AQ, v)$ and get one of dimension $n - 1 = 5$:

$$
\begin{bmatrix}
1 & -x & -1 \\
. & 1 & \frac{4}{3}y \\
. & . & 1 \\
. & . & 1 & -x \\
. & . & . & 1
\end{bmatrix} s = \begin{bmatrix}
. \\
. \\
-4 \\
3
\end{bmatrix}, \quad s = \begin{bmatrix}
3x - 4xyy \\
-4yx \\
-4y \\
x \\
3
\end{bmatrix}.
$$

The right family of the new ALS is $(1, x, 0, 1 - \frac{4}{3}xy, x - \frac{4}{3}xyx)$. After adding $4/3$-times row 5 to row 3 we can remove row 3 and column 3 and get a minimal ALS for $3x - 4xyy = (1 - \frac{4}{3}xy)3x:

$$
\begin{bmatrix}
1 & -x & -1 \\
. & 1 & \frac{4}{3}y \\
. & . & 1 & -x \\
. & . & . & 1
\end{bmatrix} s = \begin{bmatrix}
. \\
. \\
. \\
3
\end{bmatrix}, \quad s = \begin{bmatrix}
3x - 4xyy \\
-4yx \\
x \\
3
\end{bmatrix}.
$$
Factorizing a Polynomial

Now we consider the following minimal ALS for \( p = xyxyx + (3x - 4xyx) \), constructed in a similar way as shown in the previous subsection or using the right companion system (Definition 3.2):

\[
\begin{pmatrix}
1 & -x & \cdots & -x \\
. & 1 & -y & \cdots \\
. & . & 1 & -x & \frac{4}{3}x \\
. & . & . & 1 & -y \\
. & . & . & . & 1
\end{pmatrix}
\begin{pmatrix}
\alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} \\
. & 1 & \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} \\
. & . & 1 & \alpha_{3,4} & \alpha_{3,5} \\
. & . & . & 1 & \alpha_{4,5} \\
. & . & . & . & 1
\end{pmatrix}
= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\
1 & \beta_{2,3} & \beta_{2,4} & \beta_{2,5} & \beta_{2,6} \\
1 & \beta_{3,4} & \beta_{3,5} & \beta_{3,6} \\
1 & \beta_{4,5} & \beta_{4,6} \\
1 & \beta_{5,6}
\end{pmatrix}
\begin{pmatrix}
. & . & . & . & . \\
1 & -1 & -\beta_{3,4} & -\beta_{3,5} & -\beta_{3,6} \\
. & 1 & -1 & -\beta_{5,6} \\
. & . & . & . & . \\
. & . & . & . & .
\end{pmatrix}
\]

We try to create an upper right block of zeros of size 3 × 2. For that we apply the (admissible) transformation \((P, Q)\) directly to the coefficient matrices \( A_0, A_x \) and \( A_y \) in \( A = A_0 \otimes 1 + A_x \otimes x + A_y \otimes y \) to get the equations. For \( y \) we have

\[
PA_y Q = P
= \begin{pmatrix}
1 & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} \\
1 & \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} & 0 \\
1 & \alpha_{3,4} & \alpha_{3,5} & 0 & 0 \\
1 & \alpha_{4,5} & 0 & 0 & 0 \\
. & . & . & . & .
\end{pmatrix}
\begin{pmatrix}
. & . & . & . & . \\
1 & -1 & -\beta_{3,4} & -\beta_{3,5} & -\beta_{3,6} \\
. & 1 & -1 & -\beta_{5,6} \\
. & . & . & . & . \\
. & . & . & . & .
\end{pmatrix}
\]

of which we pick the upper right 3 × 2 block. Thus we have the following 6 equations for \( y \):

\[
\begin{bmatrix}
\alpha_{1,2}\beta_{3,5} + \alpha_{1,4} & \alpha_{1,4}\beta_{5,6} + \alpha_{1,2}\beta_{3,6} \\
\beta_{3,5} + \alpha_{2,4} & \alpha_{2,4}\beta_{5,6} + \beta_{3,6} \\
\alpha_{3,4} & \alpha_{3,4}\beta_{5,6}
\end{bmatrix}
= \begin{bmatrix} 0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}
\]

Similarly, we get the following 6 equations for \( x \):

\[
\begin{bmatrix}
\alpha_{1,3}\beta_{4,5} + \beta_{2,5} & \alpha_{1,3}\beta_{4,6} + \beta_{2,6} + \frac{1}{3}\alpha_{1,5} - \frac{4}{3}\alpha_{1,3} + 1 \\
\alpha_{2,3}\beta_{4,5} & \alpha_{2,3}\beta_{4,6} + \frac{1}{3}\alpha_{2,5} - \frac{4}{3}\alpha_{2,3} \\
\beta_{4,5} & \beta_{4,6} + \frac{1}{3}\alpha_{3,5} - \frac{4}{3}
\end{bmatrix}
= 0.
\]
Thus $p = (xyx - x)(yx - 3)$. The first factor is *not* atomic because we could (pre-standard admissibly) construct either an $1 \times 2$ or $2 \times 1$ upper right block of zeros in the system matrix of (4.2). On the other hand, a brief look at the ALS (4.3) immediately shows that the second factor is irreducible.
Acknowledgement

I thank Franz Lehner for the fruitful discussions and his support, Daniel Smertnig who helped me to find some orientation in abstract non-commutative factorization (especially with literature) and Roland Speicher for the creative environment in Saarbrücken.

References

[ARJ15] V. Arvind, G. Rattan, and P. Joglekar. On the complexity of noncommutative polynomial factorization. In Mathematical foundations of computer science 2015. Part II, volume 9235 of Lecture Notes in Comput. Sci., pages 38–49. Springer, Heidelberg, 2015.

[BHL17] J. P. Bell, A. Heinle, and V. Levandovskyy. On noncommutative finite factorization domains. Trans. Amer. Math. Soc., 369(4):2675–2695, 2017.

[BR11] J. Berstel and C. Reutenauer. Noncommutative rational series with applications, volume 137 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2011.

[BS15] N. R. Baeth and D. Smertnig. Factorization theory: from commutative to noncommutative settings. J. Algebra, 441:475–551, 2015.

[Car10] F. Caruso. Factorization of non-commutative polynomials. ArXiv e-prints, February 2010.

[CC80] A. Cardon and M. Crochemore. Détermination de la représentation standard d’une série reconnaissable. RAIRO Inform. Théor., 14(4):371–379, 1980.

[CLO15] D. A. Cox, J. Little, and D. O’Shea. Ideals, varieties, and algorithms. Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015. An introduction to computational algebraic geometry and commutative algebra.

[Coh63] P. M. Cohn. Noncommutative unique factorization domains. Trans. Amer. Math. Soc., 109:313–331, 1963.

[Coh72] P. M. Cohn. Generalized rational identities. In Ring theory (Proc. Conf., Park City, Utah, 1971), pages 107–115. Academic Press, New York, 1972.

[Coh85] P. M. Cohn. Free rings and their relations, volume 19 of London Mathematical Society Monographs. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, second edition, 1985.

[Coh95] P. M. Cohn. Skew fields, volume 57 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1995. Theory of general division rings.
[Coh03] P. M. Cohn. *Further algebra and applications*. Springer-Verlag London, Ltd., London, 2003.

[CR94] P. M. Cohn and C. Reutenauer. A normal form in free fields. *Canad. J. Math.*, 46(3):517–531, 1994.

[CR99] P. M. Cohn and C. Reutenauer. On the construction of the free field. *Internat. J. Algebra Comput.*, 9(3-4):307–323, 1999. Dedicated to the memory of Marcel-Paul Schützenberger.

[Fli74] M. Fliess. Matrices de Hankel. *J. Math. Pures Appl. (9)*, 53:197–222, 1974.

[FM80] E. Fornasini and G. Marchesini. On the problems of constructing minimal realizations for two-dimensional filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(2):172–176, 1980.

[Fri16] FriCAS *Computer Algebra System*, 2016. W. Hebisch, [http://axiom-wiki.newsynthesis.org/FrontPage](http://axiom-wiki.newsynthesis.org/FrontPage), svn co svn://svn.code.sf.net/p/fricas/code/trunk fricas.

[Gan86] F. R. Gantmacher. *Matrizentheorie*. Springer-Verlag, Berlin, 1986. With an appendix by V. B. Lidskij, With a preface by D. P. Želobenko, Translated from the second Russian edition by Helmut Boseck, Dietmar Soyka and Klaus Stengert.

[GHK06] A. Geroldinger and F. Halter-Koch. *Non-unique factorizations*, volume 278 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2006. Algebraic, combinatorial and analytic theory.

[GRW01] I. Gelfand, V. Retakh, and R. L. Wilson. Quadratic linear algebras associated with factorizations of noncommutative polynomials and noncommutative differential polynomials. *Selecta Math. (N.S.)*, 7(4):493–523, 2001.

[HL13] A. Heinle and V. Levandovskyy. Factorization of z-homogeneous polynomials in the first (q)-weyl algebra. *ArXiv e-prints*, January 2013.

[HV07] J. W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.*, 60(5):654–674, 2007.

[Jor89] D. A. Jordan. Unique factorisation of normal elements in noncommutative rings. *Glasgow Math. J.*, 31(1):103–113, 1989.

[Ret10] V. Retakh. From factorizations of noncommutative polynomials to combinatorial topology. *Cent. Eur. J. Math.*, 8(2):235–243, 2010.

[Sch17] K. Schrempf. Linearizing the Word Problem in (some) Free Fields. *ArXiv e-prints*, January 2017.
D. Smertnig. Factorizations of elements in noncommutative rings: A survey. *ArXiv e-prints*, July 2015.

A. Salomaa and M. Soittola. *Automata-theoretic aspects of formal power series*. Springer-Verlag, New York-Heidelberg, 1978. Texts and Monographs in Computer Science.