Linearizing the Word Problem
in (some) Free Fields

Konrad Schrempf*

January 13, 2017

Abstract

We describe a solution of the word problem in free fields (coming from non-commutative polynomials over a commutative field) using elementary linear algebra, provided that the elements are given by minimal linear representations. It relies on the normal form of Cohn and Reutenauer and can be used more generally to (positively) test rational identities. Moreover we provide a construction of minimal linear representations for the inverse of non-zero elements.

Keywords: word problem, minimal linear representation, linearization, realization, admissible linear system, rational series

AMS Classification: 16K40, 16S10, 03B25, 15A22

Introduction

Free (skew) fields arise as universal objects when it comes to embed the ring of non-commutative polynomials, that is, polynomials in (a finite number of) non-commuting variables, into a skew field [Coh85, Chapter 7]. The notion of "free fields" goes back to Amitsur [Ami66]. A brief introduction can be found in [Coh03, Section 9.3], for details we refer to [Coh95, Section 6.4]. In the present paper we restrict the setting to *commutative* ground fields, as a special case. See also [Rob84]. In [CR94], Cohn and Reutenauer introduced a *normal form* for elements in free fields in order to extend results from the theory of formal languages. In particular they characterize minimality of linear representations in terms of linear independence of

the entries of a column and a row vector, generalizing the concept of "controllability and observability matrix" (Section 3).

It is difficult to solve the word problem, that is, given linear representations (LRs for short) of two elements f and g, to decide whether $f = g$. In [CR99] the authors describe a answer of this question (for free fields coming from non-commutative polynomials over a *commutative* field). In practice however, this technique (using Gröbner bases) can be impractical even for representations of small dimensions.

Fortunately, it turns out that the word problem is equivalent to the solvability of a *linear* system of equations if *both* elements are given by *minimal* linear representations. Constructions of the latter are known for *regular* elements (non-commutative rational series), but in general non-linear techniques are necessary. This is considered in future work. Here we present a simple construction of minimal LRs for the inverses of arbitrary non-zero elements given by *minimal* LRs. In particular this applies to the inverses of non-zero polynomials with vanishing constant coefficient (which are not regular anymore).

In any case, *positive* testing of rational identities becomes easy. Furthermore, the implementation in computer (algebra) software needs only a basic data structure for matrices (linear matrix pencil) and an *exact* solver for linear systems.

Section 1 introduces the required notation concerning linear representations and admissible linear systems in free fields. Rational operations on representation level are formulated and the related concepts of linearization and realization are briefly discussed. Section 2 describes the word problem. Theorem 2.4 shows that the (in general non-linear) problem of finding appropriate transformation matrices can be reduced to a *linear system of equations* if the given LRs are *minimal*. Examples can be constructed for regular elements (rational series) as special cases (of elements in the free field), which are summarized in Section 3. Here algorithms for obtaining minimal LRs are already known. Section 4 provides a first step in the construction of minimal LRs (with linear techniques), namely for the inverses of non-zero elements given itself by a *minimal* linear representations. This is formulated in Theorem 4.20.

The main result is Theorem 2.4, the "linear" word problem. Although it is rather elementary, it opens the possibility to work *directly* on linear representations (instead of the spaces they "span"). Or, using Bergman's words [Ber78]: "The main results in this paper are trivial. But what is trivial when described in the abstract can be far from clear in the context of a complicated situation where it is needed."

1   Representing Elements

Although there are several ways for representing elements in (a subset of) the free field (linear representation [CR99], linearization [Coh85], realization [HMV06],

proper linear system [SS78], etc.) the concept of a *linear representation* seems to be the most convenient. It has the advantage (among others) that in the special case of regular elements, the general definition of the rank coincides with the Hankel rank [Fli74], [SS78, Section II.3].

Closely related to LRs are *admissible linear systems* (ALS for short) [Coh85], which could be seen as a special case. Both notations will be used synonymously. Depending on the context an ALS will be written as a triple, for example $\mathcal{A} = (u, A, v)$ or as linear system $As = v$, sometimes as $u = tA$. Like the rational operations defined on linear representation level [CR99], similar constructions can be done easily on ALS level. Thus, starting from systems for monomials (Proposition 4.1) only, a representation for *each* element in the free field can be constructed recursively.

Notation. 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 and Σ_n the permutation matrix that reverses the order of rows/columns of size n. If the size is clear from the context, I and Σ are used respectively.

Let $\mathbb{K}$ be a *commutative* field and $X = \{x_1, x_2, \ldots, x_d\}$ be a *finite* alphabet. $\mathbb{K}\langle X \rangle$ denotes the *free associative algebra* (or "algebra of non-commutative polynomials") and $\mathbb{K}(\langle X \rangle)$ denotes the *universal field of fractions* (or "free field") of $\mathbb{K}\langle X \rangle$ [Coh95], [CR99]. In the examples the alphabet is usually $X = \{x, y, z\}$.

Definition 1.1 (Inner Rank, Full Matrix, Hollow Matrix [Coh85], [CR99])**.** Given a matrix $A \in \mathbb{K}\langle X \rangle^{n \times n}$, the *inner rank* of A is the smallest number $m \in \mathbb{N}$ such that there exists a factorization $A = TU$ with $T \in \mathbb{K}\langle X \rangle^{n \times m}$ and $U \in \mathbb{K}\langle X \rangle^{m \times n}$. The matrix A is called *full* if $m = n$, *non-full* otherwise. It is called *hollow* if it contains a zero submatrix of size $k \times l$ with $k + l > n$.

Definition 1.2 (Associated and Stably Associated Matrices [Coh95])**.** Two matrices A and B over $\mathbb{K}\langle X \rangle$ (of the same size) are called *associated* over a subring $R \subseteq \mathbb{K}\langle X \rangle$ if there exist invertible matrices P, Q over R such that $A = PBQ$. A and B (not necessarily of the same size) are called *stably associated* if $A \oplus I_p$ and $B \oplus I_q$ are associated for some unit matrices I_p and I_q. Here by $C \oplus D$ we denote the diagonal sum $\begin{bmatrix} C & \cdot \\ \cdot & D \end{bmatrix}$.

In general it is hard to decide whether a matrix is full or not. For a linear matrix, that is, a matrix of the form $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$ with A_ℓ over $\mathbb{K}$, the following criterion is known, which is used in (the proof of) Theorem 2.1. If a matrix over $\mathbb{K}\langle X \rangle$ is not linear, then Higman's trick [Coh85, Section 5.8] can be used to linearize it by enlargement. The inner rank is also discussed in [FR04].

Lemma 1.3 ([Coh95, Corollary 6.3.6])**.** *A linear square matrix over $\mathbb{K}\langle X \rangle$ which is not full is associated over $\mathbb{K}$ to a linear hollow matrix.*

Definition 1.4 (Linear Representations [CR94]). Let $f \in \mathbb{K}(\!(X)\!)$. A *linear representation* of f is a triple (u, A, v) with $u \in \mathbb{K}^{1 \times n}$, $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$, $A_\ell \in \mathbb{K}^{n \times n}$ and $v \in \mathbb{K}^{n \times 1}$ such that A is full, that is, A is invertible over the free field $\mathbb{K}(\!(X)\!)$, and $f = uA^{-1}v$. The *dimension* of the representation is $\dim(u, A, v) = n$. It is called *minimal* if A has the smallest possible dimension among all linear representations of f. The "empty" representation $\pi = (,)$ is the minimal representation for $0 \in \mathbb{K}(\!(X)\!)$ with $\dim \pi = 0$.

Remark. In Definition 1.15 it can be seen that $f = uA^{-1}v$ is (up to sign) the *Schur complement* of the linearization $\left[\begin{smallmatrix} 0 & u \\ v & A \end{smallmatrix} \right]$ with respect to the upper left 1×1 block.

Definition 1.5 ([CR99]). Two linear representations are called *equivalent* if they represent the same element.

Definition 1.6 (Rank [CR99]). Let $f \in \mathbb{K}(\!(X)\!)$ and π be a *minimal* representation of f. Then the *rank* of f is defined as $\operatorname{rank} f = \dim \pi$.

Remark. The connection to the related concepts of *inversion height* and *depth* can be found in [Reu96], namely *inversion height $\leq$ depth $\leq$ rank*. Additional discussion about the depth appears in [Coh06, Section 7.7].

Definition 1.7. Let $M = M_1 \otimes x_1 + \ldots + M_d \otimes x_d$. An element in $\mathbb{K}(\!(X)\!)$ is called *regular*, if it has a linear representation (u, A, v) with $A = I - M$, that is, $A_0 = I$ in Definition 1.4, or equivalently, if A_0 is regular (invertible).

Definition 1.8 (Left and Right Families [CR94]). Let $\pi = (u, A, v)$ be a linear representation of $f \in \mathbb{K}(\!(X)\!)$ of dimension n. The families $(s_1, s_2, \ldots, s_n) \subseteq \mathbb{K}(\!(X)\!)$ with $s_i = (A^{-1}v)_i$ and $(t_1, t_2, \ldots, t_n) \subseteq \mathbb{K}(\!(X)\!)$ with $t_j = (uA^{-1})_j$ are called *left family* and *right family* respectively. $L(\pi) = \operatorname{span}\{s_1, s_2, \ldots, s_n\}$ and $R(\pi) = \operatorname{span}\{t_1, t_2, \ldots, t_n\}$ denote their linear spans.

Remark. The left family $(A^{-1}v)_i$ (respectively the right family $(uA^{-1})_j$) and the solution vector s of $As = v$ (respectively t of $u = tA$) will be used synonymously.

Proposition 1.9 ([CR94], Proposition 4.7). *A representation $\pi = (u, A, v)$ of an element $f \in \mathbb{K}(\!(X)\!)$ is minimal if and only if both, the left family and the right family are $\mathbb{K}$-linearly independent.*

Definition 1.10 (Admissible Linear Systems [Coh72]). A linear representation $\mathcal{A} = (u, A, v)$ of $f \in \mathbb{K}(\!(X)\!)$ is called *admissible linear system* (for f), denoted by $As = v$, if $u = e_1 = [1, 0, \ldots, 0]$. The element f is then the first component of the (unique) solution vector s.

Remark. In [Coh85], Cohn defines admissible linear systems with $v = v_0 \otimes 1 + v_1 \otimes x_1 + \ldots + v_d \otimes x_d$ with $v_i \in \mathbb{K}^{n \times 1}$, and $u = [0, \ldots, 0, 1]$. Writing $B = [-v, A]$ as block of size $n \times (n+1)$ the first n columns of B serve as numerator and the last

4

n columns of B as denominator. However, in this setting, for regular elements, the dimension of such a minimal system could differ from the Hankel rank [Fli74], [SS78, Section II.3].

Definition 1.11 (Admissible Transformations). Given a linear representation $\mathcal{A} = (u, A, v)$ of dimension n of $f \in \mathbb{K}(\!(X)\!)$ and invertible matrices $P, Q \in \mathbb{K}^{n \times n}$, the transformed $P\mathcal{A}Q = (uQ, PAQ, Pv)$ is again a linear representation (of f). If $\mathcal{A}$ is an ALS, the transformation (P, Q) is called *admissible* if the first row of Q is $e_1 = [1, 0, \ldots, 0]$.

Remark 1.12 (Elementary Transformations). In practice, transformations can be done by elementary row- and column operations (with respect to the system matrix A). If we add α-times row i to row $j \neq i$ in A, we also have to do this in v. If we add β-times column i to column $j \neq i$ we have to *subtract* β-times row j from row i in s. Since it is not allowed to change the first entry of s, column 1 cannot be used to eliminate entries in other columns! As an example consider the ALS

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

for the element $1 - x \in \mathbb{K}(\!(X)\!)$. Adding column 1 to column 2, that is, $Q = \begin{bmatrix} 1 & 1 \\ . & 1 \end{bmatrix}$ (and $P = I$), yields the ALS

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

for the element $-x \neq 1 - x$.

Proposition 1.13 (Rational Operations). *Let $f, g, h \in \mathbb{K}(\!(X)\!)$ be given by the admissible linear systems $\mathcal{A}_f = (u_f, A_f, v_f)$, $\mathcal{A}_g = (u_g, A_g, v_g)$ and $\mathcal{A}_h = (u_h, A_h, v_h)$ respectively, with $h \neq 0$ and let $\mu \in \mathbb{K}$. Then admissible linear systems for the rational operations can be obtained as follows:*

The scalar multiplication μf is given by

$$\mu \mathcal{A}_f = (u_f, A_f, \mu v_f).$$

The sum $f + g$ is given by

$$\mathcal{A}_f + \mathcal{A}_g = \left(\begin{bmatrix} u_f & . \end{bmatrix}, \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ . & A_g \end{bmatrix}, \begin{bmatrix} v_f \\ v_g \end{bmatrix} \right).$$

The product fg is given by

$$\mathcal{A}_f \cdot \mathcal{A}_g = \left(\begin{bmatrix} u_f & . \end{bmatrix}, \begin{bmatrix} A_f & -v_f u_g \\ . & A_g \end{bmatrix}, \begin{bmatrix} . \\ v_g \end{bmatrix} \right).$$

And the inverse h^{-1} is given by

$$\mathcal{A}_h^{-1} = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} -v_h & A_h \\ . & u_h \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right).$$

Proof. The reader can easily verify that the solution vectors for the admissible linear systems defined above are

$$\mu s_f, \quad \begin{bmatrix} s_f + u_f^\top g \\ s_g \end{bmatrix}, \quad \begin{bmatrix} s_f g \\ s_g \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} h^{-1} \\ s_h h^{-1} \end{bmatrix}$$

respectively, compare [Coh95] and [CR99]. It remains to check that the system matrices are full. For the sum and the product this is clear from the fact that the free associative algebra —being a *free ideal ring* (FIR)— has *unbounded generating number* (UGN) and therefore the diagonal sum of full matrices is full, see [Coh85, Section 7.3]. The system matrix for the inverse is full because $h \neq 0$ and therefore the linearization of $\mathcal{A}_h$ is full, compare [Coh95, Section 4.5]. □

Remark 1.14. For the rational operations from Proposition 1.13 we observe that the left families satisfy the relations

$$L(\mu \mathcal{A}_f) = L(\mathcal{A}_f),$$
$$L(\mathcal{A}_f + \mathcal{A}_g) = L(\mathcal{A}_f) + L(\mathcal{A}_g) \quad \text{and}$$
$$L(\mathcal{A}_g) \subseteq L(\mathcal{A}_f \cdot \mathcal{A}_g) = L(\mathcal{A}_f)g + L(\mathcal{A}_g).$$

And similarly for the right families we have

$$R(\mu \mathcal{A}_f) = R(\mathcal{A}_f),$$
$$R(\mathcal{A}_f + \mathcal{A}_g) = R(\mathcal{A}_f) + R(\mathcal{A}_g) \quad \text{and}$$
$$R(\mathcal{A}_f) \subseteq R(\mathcal{A}_f \cdot \mathcal{A}_g) = R(\mathcal{A}_f) + f R(\mathcal{A}_g).$$

Definition 1.15 (Linearization [BMS13], [CR99])**.** Let $f \in \mathbb{K}(\!(X)\!)$. A *linearization* of f is a matrix $L = L_0 \otimes 1 + L_1 \otimes x_1 + \ldots + L_d \otimes x_d$, with $L_\ell \in \mathbb{K}^{m \times m}$, of the form

$$L = \begin{bmatrix} c & u \\ v & A \end{bmatrix} \quad \in \mathbb{K}\langle X \rangle^{m \times m}$$

such that A is invertible over the free field and f is the Schur complement, that is, $f = c - uA^{-1}v$. If $c = 0$ then L is called a *pure* linearization. The *size* of the linearization is size $L = m$, the *dimension* is dim $L = m - 1$.

Proposition 1.16 ([BMS13, Proposition 3.2])**.** *Let $\mathbb{F} = \mathbb{K}(\!(X)\!)$ and $A \in \mathbb{F}^{k \times k}$, $B \in \mathbb{F}^{k \times l}$, $C \in \mathbb{F}^{l \times k}$ and $D \in \mathbb{F}^{l \times l}$ be given and assume that D is invertible in*

$\mathbb{F}^{l \times l}$. Then the matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is invertible in $\mathbb{F}^{(k+l) \times (k+l)}$ if and only if the Schur complement $A - BD^{-1}C$ is invertible in $\mathbb{F}$. In this case

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} . & . \\ . & D^{-1} \end{bmatrix} + \begin{bmatrix} I_k \\ -D^{-1}C \end{bmatrix} \left(A - BD^{-1}C\right)^{-1} \begin{bmatrix} I_k & -BD^{-1} \end{bmatrix}.$$

Remark 1.17. (i) Let $f \in \mathbb{K}(\!(X)\!)$ be given by the linearization L. Then $f = |L|_{1,1}$ is the $(1,1)$-quasideterminant [GGRW05] of L.

(ii) Given a linear representation (u, A, v) of $f \in \mathbb{K}(\!(X)\!)$, then $L = \begin{bmatrix} . & u \\ -v & A \end{bmatrix}$ is a pure linearization of f.

(iii) Talking about a *minimal* linearization, one has to specify which class of matrices is considered: Scalar entries in the first row and column? Pure? And, if applicable, selfadjoint?

Proposition 1.18. *Let*

$$L = \begin{bmatrix} c & u \\ v & A \end{bmatrix}$$

be a linearization of size n for some element $f \in \mathbb{K}(\!(X)\!)$ and define another element $g \in \mathbb{K}(\!(X)\!)$ by the pure linearization

$$\tilde{L} = \begin{bmatrix} . & \tilde{u} \\ \tilde{v} & \tilde{A} \end{bmatrix} \quad \text{with} \quad \tilde{A} = \begin{bmatrix} c & u & -1 \\ v & A & . \\ -1 & . & . \end{bmatrix}, \quad \tilde{u} = [0, \ldots, 0, 1], \quad \tilde{v} = \tilde{u}^{\top}$$

of size $n + 2$. Then $g = f$.

Proof. Using Proposition 1.16 —taking the Schur complement with respect to the block entry $(2,2)$— and $b = [-1, 0, \ldots, 0]$, the inverse of $\tilde{A} = \begin{bmatrix} L & b^{\top} \\ b & . \end{bmatrix}$ can be written as

$$\tilde{A}^{-1} = \begin{bmatrix} L^{-1} & . \\ . & . \end{bmatrix} - \begin{bmatrix} -L^{-1}b^{\top} \\ 1 \end{bmatrix} \left(bL^{-1}b^{\top}\right)^{-1} \begin{bmatrix} -bL^{-1} & 1 \end{bmatrix}.$$

Hence

$$-\tilde{u}\tilde{A}^{-1}\tilde{v} = -\left(\begin{bmatrix} . & . \end{bmatrix} - \left(bL^{-1}b^{\top}\right)^{-1} \begin{bmatrix} -bL^{-1} & 1 \end{bmatrix}\right) \begin{bmatrix} . \\ 1 \end{bmatrix}$$

$$= \left(bL^{-1}b^{\top}\right)^{-1}$$

$$= \left(b\left(\begin{bmatrix} . & . \\ . & A^{-1} \end{bmatrix} + \begin{bmatrix} 1 \\ -A^{-1}v \end{bmatrix} \left(c - uA^{-1}v\right)^{-1} \begin{bmatrix} 1 & -uA^{-1} \end{bmatrix}\right)b^{\top}\right)^{-1}$$

$$= \left(\left(\begin{bmatrix} . & . \end{bmatrix} - \left(c - uA^{-1}v\right)^{-1} \begin{bmatrix} 1 & -uA^{-1} \end{bmatrix}\right) \begin{bmatrix} -1 \\ . \end{bmatrix}\right)^{-1}$$

$$= c - uA^{-1}v.$$

7

□

If the first row and column of a linearization for some $f \in \mathbb{K}(\!(X)\!)$ contains non-scalar entries, then Proposition 1.18 can be used to construct a linear representation of f. On the other hand, given a linear representation of dimension n (of f) which can be brought to such a form, a linearization of size $n - 1$ can be obtained. The characterization of minimality for linearizations will be considered in future work.

Example 1.19. For the anticommutator $xy + yx$ a minimal ALS is given by

$$
\left(\begin{bmatrix} 1 & . & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & -y & . \\ . & 1 & . & -y \\ . & . & 1 & -x \\ . & . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix} \right).
$$

Permuting the columns and multiplying the system matrix by -1 we get the linearization

$$
L'_{xy+yx} = \begin{bmatrix} . & . & . & . & 1 \\ . & . & y & x & -1 \\ . & y & . & -1 & . \\ . & x & -1 & . & . \\ 1 & -1 & . & . & . \end{bmatrix}
$$

which is of the form in Proposition 1.18 and yields a *minimal* (pure) linearization of the anticommutator

$$
L_{xy+yx} = \begin{bmatrix} . & y & x \\ y & . & -1 \\ x & -1 & . \end{bmatrix}.
$$

Definition 1.20 (Realization [HMV06]). A *realization* of a matrix $F \in \mathbb{K}(\!(X)\!)^{p \times q}$ is a quadruple (A, B, C, D) with $A = A_0 \otimes 1 + A_1 \otimes x_1 + \ldots + A_d \otimes x_d$, $A_\ell \in \mathbb{K}^{n \times n}$, $B \in \mathbb{K}^{n \times q}$, $C \in \mathbb{K}^{p \times n}$ and $D \in \mathbb{K}^{p \times q}$ such that A is invertible over the free field and $F = D - CA^{-1}B$. The *dimension* of the realization is $\dim(A, B, C, D) = n$.

Remark. A realization $\mathcal{R} = (A, B, C, D)$ could be written in block form

$$
L_{\mathcal{R}} = \begin{bmatrix} D & C \\ B & A \end{bmatrix} \quad \in \mathbb{K}\langle X \rangle^{(p+n) \times (q+n)}.
$$

Here, the definition is such that $F = |L_{\mathcal{R}}|_{1',1'}$ is the $(1,1)$-block-quasideterminant [GGRW05] with respect to block D. For $A = -J + L_A(X)$ we obtain the *descriptor realization* in [HMV06]. Realizations where B and/or C contain non-scalar entries are sometimes called "butterfly realizations" [HMV06]. Minimality with respect to realizations is investigated in [Vol15].

8

2 The Word Problem

Let $f, g \in \mathbb{K}(\langle X \rangle)$ be given by the linear representations $\pi_f = (u_f, A_f, v_f)$ and $\pi_g = (u_g, A_g, v_g)$ of dimension n_f and n_g respectively and define the matrix

$$M = \begin{bmatrix} . & u_f & u_g \\ v_f & A_f & . \\ v_g & . & -A_g \end{bmatrix},$$

which is a linearization of $f - g$, of size $n = n_f + n_g + 1$. Then $f = g$ if and only if M is not full [Coh95, Section 4.5]. For the word problem see also [Coh95, Section 6.6]. Whether M is full or not can be decided by the following theorem.

Theorem 2.1 ([CR99, Theorem 4.1]). *For each $r \in \{1, 2, \ldots, n\}$, denote by I_r the ideal of $\mathbb{K}[a, b]$ generated by the polynomials $\det A - 1$, $\det B - 1$ and the coefficients of each $x \in \{1\} \cup X$ in the (i, j) entries of the matrix AMB for $1 \leq i \leq r$, $r \leq j \leq n$. Then the linear matrix M is full if and only if for each $r \in \{1, 2, \ldots, n\}$, the ideal I_r is trivial.*

Remark. Notice that there is a misprint in [CR99] and the coefficients of M_0 are omitted.

So far we were not able to apply this theorem practically for $n \geq 5$, where 50 or more unknowns are involved. However, if we have any ALS (or linear representation) for $f - g$, say from Proposition 1.13, then we could check whether it can be (admissibly) transformed into a smaller system, for example $A's' = 0$. For polynomials (with $A = I - Q$ and Q upper triangular and nilpotent) this could be done row by row. In general the pivot blocks (the blocks in the diagonal) can be arbitrarily large. Therefore this elimination has to be done *blockwise* by setting up a single linear system for row *and* column operations. This idea is used in the following lemma. Note that the existence of a solution for this linear system is *invariant* under admissible transformations (on the subsystems). This is a key requirement since the normal form [CR94] is defined modulo similarity transformations (more general by stable association, Definition 1.2).

Theorem 2.2 ([CR99, Theorem 1.4]). *If $\pi' = (u', A', v')$ and $\pi'' = (u'', A'', v'')$ are equivalent (pure) linear representations, of which the first is minimal, then the second is isomorphic to a representation (u, A, v) which has the block decomposition*

$$u = \begin{bmatrix} * & u' & . \end{bmatrix}, \quad A = \begin{bmatrix} * & . & . \\ * & A' & . \\ * & * & * \end{bmatrix} \quad and \quad v = \begin{bmatrix} . \\ v' \\ * \end{bmatrix}.$$

Lemma 2.3. *Let $f, g \in \mathbb{K}(\!(X)\!)$ be given by the admissible linear systems $\mathcal{A}_f = (u_f, A_f, v_f)$ and $\mathcal{A}_g = (u_g, A_g, v_g)$ of dimension n_f and n_g respectively. If there exist matrices $T, U \in \mathbb{K}^{n_f \times n_g}$ such that $u_f U = 0$, $T A_g - A_f U = A_f u_f^\top u_g$ and $T v_g = v_f$, then $f = g$.*

Proof. The difference $f - g$ can be represented by the admissible linear system $As = v$ with

$$A = \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ \cdot & A_g \end{bmatrix}, \quad s = \begin{bmatrix} s_f - u_f^\top g \\ -s_g \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_f \\ -v_g \end{bmatrix}.$$

Defining the (invertible) transformations

$$P = \begin{bmatrix} I_{n_f} & T \\ \cdot & I_{n_g} \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} I_{n_f} & -U \\ \cdot & I_{n_g} \end{bmatrix}$$

and $A' = PAQ$, $s' = Q^{-1}s$ and $v' = Pv$ we get a new system $A's' = v'$:

$$\begin{aligned} A' &= \begin{bmatrix} I_{n_f} & T \\ \cdot & I_{n_g} \end{bmatrix} \begin{bmatrix} A_f & -A_f u_f^\top u_g \\ \cdot & A_g \end{bmatrix} \begin{bmatrix} I_{n_f} & -U \\ \cdot & I_{n_g} \end{bmatrix} \\ &= \begin{bmatrix} A_f & -A_f u_f^\top u_g + T A_g \\ \cdot & A_g \end{bmatrix} \begin{bmatrix} I_{n_f} & -U \\ \cdot & I_{n_g} \end{bmatrix} \\ &= \begin{bmatrix} A_f & -A_f u_f^\top u_g + T A_g - A_f U \\ \cdot & A_g \end{bmatrix} = \begin{bmatrix} A_f & 0 \\ \cdot & A_g \end{bmatrix}, \\ s' &= \begin{bmatrix} I_{n_f} & U \\ \cdot & I_{n_g} \end{bmatrix} \begin{bmatrix} s_f - u_f^\top g \\ -s_g \end{bmatrix} = \begin{bmatrix} s_f - u_f^\top g - U s_g \\ -s_g \end{bmatrix}, \\ v' &= \begin{bmatrix} I_{n_f} & T \\ \cdot & I_{n_g} \end{bmatrix} \begin{bmatrix} v_f \\ -v_g \end{bmatrix} = \begin{bmatrix} v_f - T v_g \\ -v_g \end{bmatrix}. \end{aligned}$$

Invertibility of A' over the free field implies $s_f - u_f^\top g - U s_g = 0$, in particular

$$\begin{aligned} 0 &= u_f s_f - u_f u_f^\top g - u_f U s_g \\ &= f - g \end{aligned}$$

because $u_f U = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Let d be the number of letters in the alphabet X, $\dim \mathcal{A}_f = n_f$ and $\dim \mathcal{A}_g = n_g$. To determine the transformation matrices $T, U \in \mathbb{K}^{n_f \times n_g}$ from the lemma we just have to solve a linear system of $(d+1)n_f(n_g+1)$ equations in $2n_f n_g$ unknowns. If there is a solution then $f = g$. Neither $\mathcal{A}_f$ nor $\mathcal{A}_g$ have to be minimal. Computer experiments show, that Hua's identity [Ami66]

$$x - \left(x^{-1} + (y^{-1} - x)^{-1}\right)^{-1} = xyx$$

10

can be tested positively by Lemma 2.3 when the left hand side is constructed by the rational operations from Proposition 1.13. However, without assuming minimality, the fact that there is no solution does *not* imply, that $f \neq g$, see Example 2.5 below.

Theorem 2.4 ("Linear solution" of the Word Problem). *Let $f, g \in \mathbb{K}(\!(X)\!)$ be given by the* minimal *admissible linear systems $\mathcal{A}_f = (u_f, A_f, v_f)$ and $\mathcal{A}_g = (u_g, A_g, v_g)$ of dimension n respectively. Then $f = g$ if and only if there exist matrices $T, U \in \mathbb{K}^{n \times n}$ such that $u_f U = 0$, $T A_g - A_f U = A_f u_f^\top u_g$ and $T v_g = v_f$.*

Proof. If $f = g$ then, since admissible linear systems correspond to (pure) linear representations, by Theorem 2.2 there exist invertible matrices $P, Q \in \mathbb{K}^{n \times n}$ such that $A_f = P A_g Q$ and $v_f = P v_g$. Let $T = P$ and $U = Q^{-1} - u_f^\top u_g$. The admissible linear systems are minimal. Hence, the left family s_f is $\mathbb{K}$-linearly independent. Since the first component of s_g is equal to the first component of $s_f = Q^{-1} s_g$ and the left family s_g is $\mathbb{K}$-linearly independent, the first row of Q^{-1} must be $[1, 0, \ldots, 0]$. Therefore $u_f U = u_f (Q^{-1} - u_f^\top u_g) = 0$. Clearly $v_f = T v_g$ and

$$
\begin{aligned}
T A_g - A_f U &= P A_g - A_f Q^{-1} + A_f u_f^\top u_g \\
&= A_f u_f^\top u_g.
\end{aligned}
$$

The other implication follows from Lemma 2.3. □

Example 2.5. Let $f = x^{-1}$ and $g = x^{-1}$ be given by the admissible linear systems

$$
[x] s_f = [1] \quad \text{and} \quad \begin{bmatrix} x & -z \\ . & 1 \end{bmatrix} s_g = \begin{bmatrix} 1 \\ . \end{bmatrix}
$$

respectively. Then the ALS

$$
\begin{bmatrix} x & -x & . \\ . & x & -z \\ . & . & 1 \end{bmatrix} s = \begin{bmatrix} 1 \\ -1 \\ . \end{bmatrix}, \quad s = \begin{bmatrix} 0 \\ -x^{-1} \\ 0 \end{bmatrix}
$$

represents $f - g = 0$.

While it is obvious here that the second component of the solution vector s_g is zero, it is not clear in general how one can exclude such "pathological" LRs without minimality assumption. One might ask, for which class of constructions (rational operations [CR99], Higman's trick [Hig40, Coh85], selfadjoint linearization trick [And13], etc.) there are sufficient conditions for the existence of matrices T, U (over $\mathbb{K}$) in Lemma 2.3 if $f = g$. Unfortunately this seems to be impossible except for some specific examples as the following ALS (constructed by the rational operations from Proposition 1.13) for $x - xyy^{-1} = 0$ suggests (some zeros are kept to emphasize

the block structure):

$$
\begin{bmatrix}
1 & -x & -1 & . & . & . & . & . & . \\
0 & 1 & . & . & . & . & . & . & . \\
. & . & 1 & -x & . & . & . & . & . \\
. & . & 0 & 1 & -1 & . & . & . & . \\
. & . & . & . & 0 & 1 & -y & . & . \\
. & . & . & . & -1 & 0 & 1 & . & . \\
. & . & . & . & 0 & 1 & 0 & -1 & . \\
. & . & . & . & . & . & . & 1 & -y \\
. & . & . & . & . & . & . & 0 & 1
\end{bmatrix}
\quad
s =
\begin{bmatrix}
. \\ 1 \\ . \\ . \\ . \\ . \\ . \\ . \\ -1
\end{bmatrix} .
$$

There are no T, U and therefore no P, Q (admissible, with blocks T, U) such that PAQ has a 2×7 upper right block of zeros and the first two components of Pv are zero. Therefore we would like to construct minimal linear representations directly. For regular elements, algorithms are known, see Section 3. How to proceed in general is open except for the inverse. This is discussed in Section 4.

3   Regular Elements

For regular elements (Definition 1.7) in the free field minimal linear representations can be obtained via the Extended Ho-Algorithm [FM80] from the Hankel matrix or by minimizing a given linear representation via the algorithm in [CC80] by detecting linearly dependent rows in the controllability matrix and linearly dependent columns in the observability matrix, see Definition 3.3. The basic idea goes back to Schützenberger [Sch61]. Controllability and observability is discussed in [KFA69, Chapter 10].

For an alphabet $X = \{x_1, x_2, \ldots, x_d\}$ (finite, non-empty set), the free monoid generated by X is denoted by X^*. A formal power series (in non-commuting variables) is a mapping f from X^* to a commutative field $\mathbb{K}$, written as formal sum

$$
f = \sum_{w \in X^*} (f, w) \, w
$$

with coefficients $(f, w) \in \mathbb{K}$. In general, $\mathbb{K}$ could be replaced by a ring or a skew field. [SS78] and [BR11] contain detailed introductions. On the set of formal power series $\mathbb{K}\langle\!\langle X \rangle\!\rangle$ the following rational operations are defined for $f, g, h \in \mathbb{K}\langle\!\langle X \rangle\!\rangle$ with $(h, 1) = 0$, and $\mu \in \mathbb{K}$:

The scalar multiplication

$$
\mu f = \sum_{w \in X^*} \mu(f, w) \, w.
$$

The sum

$$f + g = \sum_{w \in X^*} \big((f, w) + (g, w)\big)\, w.$$

The product

$$fg = \sum_{w \in X^*} \left(\sum_{uv=w} (f, u)(g, v) \right) w.$$

And the quasiinverse

$$h^+ = \sum_{n \geq 1} h^n.$$

The set of non-commutative (nc) rational series $\mathbb{K}^{\mathrm{rat}} \langle\!\langle X \rangle\!\rangle$ is the smallest rationally closed (that is, closed under scalar multiplication, sum, product and quasiinverse) subset of $\mathbb{K} \langle\!\langle X \rangle\!\rangle$ containing the nc polynomials $\mathbb{K}\langle X \rangle$. A series $f \in \mathbb{K} \langle\!\langle X \rangle\!\rangle$ is called *recognizable* if there exists a natural number n, a monoid homomorphism $\mu : X^* \to \mathbb{K}^{n \times n}$ and two vectors $\alpha \in \mathbb{K}^{1 \times n}$, $\beta \in \mathbb{K}^{n \times 1}$ such that f can be written as

$$f = \sum_{w \in X^*} \alpha\, \mu(w)\, \beta\, w.$$

The triple (α, μ, β) is called a *linear representation* [SS78, Section II.2].

Theorem 3.1 ([Sch61]). *A series $f \in \mathbb{K} \langle\!\langle X \rangle\!\rangle$ is rational if and only if it is recognizable.*

A rational series f can be represented by a *proper linear system* (PLS for short) $s = v + Qs$ where f is the first component of the *unique* solution vector s (with $v \in \mathbb{K}^{n \times 1}$, $Q = Q_1 \otimes x_1 + Q_2 \otimes x_2 + \ldots + Q_d \otimes x_d$, $Q_i \in \mathbb{K}^{n \times n}$ for some $n \in \mathbb{N}$). Rational operations are then formulated on this level [SS78, Section II.1]. Clearly, every proper linear system gives rise to an admissible linear system $\mathcal{A} = (u, I - Q, v)$ with $u = e_1$. When (α, μ, β) is a linear representation of a recognizable series, then $\pi = (\alpha, I - Q, \beta)$ with

$$Q = \sum_{x \in X} \mu(x) \otimes x. \tag{3.2}$$

is a linear representation of $f \in \mathbb{K} \langle\!\langle X \rangle\!\rangle$. For a PLS the solution vector s can be computed by the quasiinverse Q^+:

$$\begin{aligned} s &= (I - Q)^{-1} v \\ &= (I + Q^+) v \\ &= (I + Q + Q^2 + Q^3 + \ldots) v. \end{aligned}$$

Definition 3.3 (Controllability and Observability Matrix). Let $\mathcal{P} = (u, I - Q, v)$ be a proper linear system of dimension n (for some nc rational series). Then the *controllability matrix* and the *observability matrix* are defined as

$$\mathcal{V} = \mathcal{V}(\mathcal{P}) = \begin{bmatrix} v & Qv & \cdots & Q^{n-1}v \end{bmatrix} \quad \text{and} \quad \mathcal{U} = \mathcal{U}(\mathcal{P}) = \begin{bmatrix} u \\ uQ \\ \vdots \\ uQ^{n-1} \end{bmatrix}$$

respectively.

Remark. Note that the monomials (in the polynomials) in Q^k have length k. The matrices $\mathcal{V}$ and $\mathcal{U}$ are over $\mathbb{K}\langle X \rangle$. A priori these matrices would have an infinite number of columns and rows respectively. However, by [Coh95, Lemma 6.6.3], it suffices to use the columns of $\mathcal{V}$ and rows of $\mathcal{U}$ only. This gives the connection to [BR11, Section I.2] and could be used for minimizing proper linear systems [SS78]. In other words: Instead of identifying $\mathbb{K}$-linear dependence of the left family $s = (I - Q)^{-1}v = (I + Q + Q^2 + \ldots)v$, we can restrict to the "approximated" left family $\tilde{s} = (I + Q + \ldots + Q^{n-1})v$.

Now let $X_k^* \subseteq X^*$ denote the set of words of length k and use $\mu : X^* \to \mathbb{K}^{n \times n}$ from (3.2) to define $V_k \in \mathbb{K}^{n \times d^k}$ with columns $\mu(w)v$ for $w \in X_k^*$ and $U_k \in \mathbb{K}^{d^k \times n}$ with rows $u\mu(w)$ for $w \in X_k^*$. Then the *controllability matrix* and the *observability matrix* can be defined alternatively as

$$\mathcal{V}' = \begin{bmatrix} V_0 & V_1 & \cdots & V_{n-1} \end{bmatrix} \quad \text{and} \quad \mathcal{U}' = \begin{bmatrix} U_0 \\ U_1 \\ \vdots \\ U_{n-1} \end{bmatrix} \quad \text{respectively}$$

with entries in $\mathbb{K}$. Note that the rank of $\mathcal{V}'$ is at most n while the number of columns of the blocks V_k is d^k. So —for an alphabet with more than one letter— most of the columns are not needed. For $X = \{x\}$ and $Q = Q_x \otimes x$ we have $V_k = Q_x^k v$ and $U_k = uQ_x^k$. Hence $\mathcal{V}'$ and $\mathcal{U}'$ can be written as

$$\mathcal{V}' = \begin{bmatrix} v & Q_x v & \cdots & Q_x^{n-1}v \end{bmatrix} \quad \text{and} \quad \mathcal{U}' = \begin{bmatrix} u \\ uQ_x \\ \vdots \\ uQ_x^{n-1} \end{bmatrix} \quad \text{respectively.}$$

Compare with [KFA69, Section 6.3]. For controllability and observability in connection with realizations (Definition 1.20) see also [HMV06].

14

4 Minimizing the Inverse

Having solved the word problem, our next goal is "minimal arithmetics" in the free field $\mathbb{K}(\!(X)\!)$. That is, given elements by minimal admissible linear systems, to compute minimal ones for the rational operations. For the scalar multiplication this is trivial. For the inverse some preparation is necessary. The result is presented in Theorem 4.20. The "minimal sum" and the "minimal product" are considered in future work. The main difficulty is *not* minimality but the restriction to *linear* techniques.

Proposition 4.1. *Let* $k \in \mathbb{N}$ *and* $f = x_{i_1} x_{i_2} \cdots x_{i_k}$ *be a monomial in* $\mathbb{K}\langle X \rangle \subseteq \mathbb{K}(\!(X)\!)$. *Then*

$$
\mathcal{A} = \left(\begin{bmatrix} 1 & . & \cdots & . \end{bmatrix}, \begin{bmatrix} 1 & -x_{i_1} & & & \\ & 1 & -x_{i_2} & & \\ & & \ddots & \ddots & \\ & & & 1 & -x_{i_k} \\ & & & & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \vdots \\ . \\ 1 \end{bmatrix} \right)
$$

is a minimal ALS of dimension $\dim \mathcal{A} = k + 1$.

Proof. The system matrix of $\mathcal{A}$ is full. For row indices $[1, x_{i_1}, x_{i_1} x_{i_2}, \ldots, x_{i_1} \cdots x_{i_k}]$ and column indices $[1, x_{i_k}, x_{i_{k-1}} x_{i_k}, \ldots, x_{i_1} \cdots x_{i_k}]$ the Hankel matrix [Fli74], [SS78, Section II.3] of f is

$$
H(f) = \begin{bmatrix} & & 1 \\ & 1 & \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}
$$

and has rank $k + 1$. $\qquad\square$

Remark. Trivially, $\mathcal{A} = ([1], [1], [1])$ is a minimal ALS for the *unit element* (empty word).

The following proposition is a variant of the inverse in Proposition 1.13 and is motivated by inverting the inverse of a monomial, for example, $f = (xyz)^{-1}$. A minimal ALS for f is given by

$$
\begin{bmatrix} z & -1 & . \\ . & y & -1 \\ . & . & x \end{bmatrix} s = \begin{bmatrix} . \\ . \\ 1 \end{bmatrix}, \qquad s = \begin{bmatrix} z^{-1} y^{-1} x^{-1} \\ y^{-1} x^{-1} \\ x^{-1} \end{bmatrix}.
$$

Minimality is clear immediately by also checking the $\mathbb{K}$-linear independence of the right family. Using the construction of the inverse from Proposition 1.13 we get the

15

system

$$\begin{bmatrix} . & z & -1 & . \\ . & . & y & -1 \\ -1 & . & . & x \\ . & 1 & . & . \end{bmatrix} s = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}, \qquad s = \begin{bmatrix} xyz \\ 1 \\ z \\ yz \end{bmatrix}.$$

for $f^{-1} = xyz$. To obtain the form of Proposition 4.1 we have to reverse the rows $1, 2, 3$ and the columns $2, 3, 4$ and multiply the rows $1, 2, 3$ by -1.

Proposition 4.2 (Standard Inverse). *Let $0 \neq f \in \mathbb{K}(\langle X \rangle)$ be given by the admissible linear system $\mathcal{A} = (u, A, v)$ of dimension n. Then an admissible linear system of dimension $n + 1$ for f^{-1} is given by*

$$\mathcal{A}^{-1} = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} \Sigma v & -\Sigma A \Sigma \\ . & u \Sigma \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right). \tag{4.3}$$

(Recall that the permutation matrix $\Sigma = \Sigma_n$ reverses the order of rows/columns.)

Definition 4.4 (Standard Inverse). Let $\mathcal{A}$ be an ALS for a non-zero element. We call the ALS (4.3) the *standard inverse* of $\mathcal{A}$, denoted by $\mathcal{A}^{-1}$.

Proof. The reader can easily verify that the solution vector of $\mathcal{A}^{-1}$ is

$$\begin{bmatrix} f^{-1} \\ \Sigma_n s_f f^{-1} \end{bmatrix}.$$

Compare with Proposition 1.13. $\qquad\square$

We proceed with the calculation of minimal admissible linear systems for the inverse. We distinguish four types of ALS according to the form of the system matrix. Later, in the remark following Lemma 4.18 and 4.19, we will see how to bring a system matrix to one of these forms depending on the left and right families.

Lemma 4.5 (Inverse Type $(1,1)$). *Assume that $0 \neq f \in \mathbb{K}(\langle X \rangle)$ has a minimal admissible linear system of dimension n of the form*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right). \tag{4.6}$$

Then a minimal ALS for f^{-1} of dimension $n - 1$ is given by

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . \end{bmatrix}, \begin{bmatrix} -\lambda \Sigma b'' & -\Sigma B \Sigma \\ -\lambda b & -b' \Sigma \end{bmatrix}, \begin{bmatrix} . \\ 1 \end{bmatrix} \right) \tag{4.7}$$

with $1 \notin R(\mathcal{A}')$ and $1 \notin L(\mathcal{A}')$.

Proof. The standard inverse of $\mathcal{A}$ is

$$
\begin{bmatrix}
\lambda & -1 & . & . \\
. & -\Sigma b'' & -\Sigma B\Sigma & . \\
. & -b & -b'\Sigma & -1 \\
. & . & . & 1
\end{bmatrix}
\tilde{s} =
\begin{bmatrix}
. \\
. \\
. \\
1
\end{bmatrix}.
$$

Adding row 4 to row 3 and λ-times column 2 to column 1 gives

$$
\begin{bmatrix}
0 & -1 & . & . \\
-\lambda\Sigma b'' & -\Sigma b'' & -\Sigma B\Sigma & . \\
-\lambda b & -b & -b'\Sigma & 0 \\
. & . & . & 1
\end{bmatrix}
s' =
\begin{bmatrix}
. \\
. \\
1 \\
1
\end{bmatrix}.
$$

It follows that $s_2' = 0$ and s_{n+1}' does not contribute to the solution s_1' and thus the first and the last row as well as the second and the last column can be removed. If $\mathcal{A}'$ were not minimal, then there would exist a system $\mathcal{A}''$ of dimension $m < n - 1$ for f^{-1}. The standard inverse $(\mathcal{A}'')^{-1}$ would give a system of dimension $m + 1 < n$ for f, contradicting minimality of $\mathcal{A}$. It remains to show that $1 \notin R(\mathcal{A}')$ and $1 \notin L(\mathcal{A}')$. Let $t = (t_1, t_2, \ldots, t_n)$ be the right family of $\mathcal{A}$ which is (due to minimality) $\mathbb{K}$-linearly independent. Then the right family of $\mathcal{A}^{-1}$ is $(t_n f^{-1}, \ldots, t_2 f^{-1}, t_1 f^{-1}, f^{-1})$, that after the row operation becomes $(t_n, \ldots, t_2, t_1, 1 - t_1) f^{-1}$. Removing the first and the last component (corresponding to the first and the last row) yields the right family $(t_{n-1}, \ldots, t_2, t_1) f^{-1}$. Therefore $1 \notin R(\mathcal{A}')$, because otherwise $f \in \mathrm{span}\{t_{n-1}, \ldots, t_2, t_1\}$, contradicting $\mathbb{K}$-linear independence of t. Similar arguments show that $1 \notin L(\mathcal{A}')$. $\square$

Lemma 4.8 (Inverse Type $(1,0)$)**.** *Assume that* $0 \neq f \in \mathbb{K}\langle\!\langle X \rangle\!\rangle$ *has a minimal admissible linear system of dimension n of the form*

$$
\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & b' & b \\ . & B & b'' \\ . & c' & c \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right)
\tag{4.9}
$$

with $1 \notin L(\mathcal{A})$. Then a minimal ALS for f^{-1} of dimension n is given by

$$
\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -\frac{1}{\lambda}c & -\frac{1}{\lambda}c'\Sigma \\ . & -\Sigma b'' & -\Sigma B\Sigma \\ . & -b & -b'\Sigma \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)
\tag{4.10}
$$

with $1 \notin L(\mathcal{A}')$.

Proof. The standard inverse of $\mathcal{A}$ is

$$
\begin{bmatrix}
\lambda & -c & -c'\Sigma & . \\
. & -\Sigma b'' & -\Sigma B\Sigma & . \\
. & -b & -b'\Sigma & -1 \\
. & . & . & 1
\end{bmatrix}
\tilde{s} =
\begin{bmatrix}
. \\
. \\
. \\
1
\end{bmatrix}
$$

and has dimension $n + 1$. After adding row $n + 1$ to row n we can remove row and column $n + 1$ because $\tilde{s}_{n+1}$ does not contribute to the solution $\tilde{s}_1 = f^{-1}$. Then we divide the first row by λ and obtain (4.10). It remains to show that $\mathcal{A}'$ is minimal and $1 \notin L(\mathcal{A}')$. Let $(s_1, s_2, \ldots, s_n)$ be the left family of $\mathcal{A}$ which is (due to minimality) $\mathbb{K}$-linearly independent. Then the left family of $\mathcal{A}^{-1}$ is $(f^{-1}, s_n f^{-1}, \ldots, s_2 f^{-1}, 1)$. Note that (admissible) row operations do not affect the left family. Since we removed the last entry $s_1 f^{-1} = \tilde{s}_{n+1} = 1$, the left family of $\mathcal{A}'$ is $(1, s_n, \ldots, s_2) f^{-1}$. By assumption $1 \notin L(\mathcal{A})$. Therefore $1 \notin \operatorname{span}\{s_2, s_3, \ldots, s_n\}$, hence $(1, s_n, \ldots, s_2)$ is $\mathbb{K}$-linearly independent. Clearly, $1 \notin L(\mathcal{A}')$ because $f \notin \operatorname{span}\{1, s_n, \ldots, s_2\}$. Similarly, let $(t_1, t_2, \ldots, t_n)$ be the right family of $\mathcal{A}$ which is $\mathbb{K}$-linearly independent. Then the right family of $\mathcal{A}^{-1}$ is $(t_n f^{-1}, \ldots, t_2 f^{-1}, t_1 f^{-1}, f^{-1})$, that after the row operation is $(t_n f^{-1}, \ldots, t_2 f^{-1}, t_1 f^{-1}, f^{-1} - t_1 f^{-1})$. Since we removed the last entry, the right family of $\mathcal{A}'$ is $(t_n, \ldots, t_2, t_1) f^{-1}$ which is clearly $\mathbb{K}$-linearly independent. Proposition 1.9 gives minimality of $\mathcal{A}'$. □

Lemma 4.11 (Inverse Type $(0,1)$). *Assume that $0 \neq f \in \mathbb{K}(\!(X)\!)$ has a minimal admissible linear system of dimension n of the form*

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} a & b' & b \\ a' & B & b'' \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ \lambda \end{bmatrix} \right) \tag{4.12}$$

with $1 \notin R(\mathcal{A})$. Then a minimal ALS for f^{-1} of dimension n is given by

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} -\lambda \Sigma b'' & -\Sigma B \Sigma & -\Sigma a'. \\ -\lambda b & -b' \Sigma & -a \\ . & . & 1 \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right) \tag{4.13}$$

with $1 \notin R(\mathcal{A}')$.

Proof. The standard inverse of $\mathcal{A}$ is

$$\begin{bmatrix} \lambda & -1 & . & . \\ . & -\Sigma b'' & -\Sigma B \Sigma & -\Sigma a' \\ . & -b & -b' \Sigma & -a \\ . & . & . & 1 \end{bmatrix} \tilde{s} = \begin{bmatrix} . \\ . \\ . \\ 1 \end{bmatrix}.$$

After adding λ-times column 2 to column 1 we can remove row 1 and column 2, because $\tilde{s}_2 = 0$. Showing minimality and $1 \notin R(\mathcal{A}')$ is similar to the proof of Lemma 4.8 (column operations affect the left family). □

Lemma 4.14 (Inverse Type $(0,0)$). *Let $\mathcal{A} = (u, A, v)$ be a minimal admissible linear system of dimension n for $0 \neq f \in \mathbb{K}(\!(X)\!)$ with $1 \notin R(\mathcal{A})$ and $1 \notin L(\mathcal{A})$. Then the standard inverse $\mathcal{A}^{-1}$ is a minimal ALS of dimension $n + 1$ for f^{-1}.*

Proof. If $\mathcal{A}^{-1}$ were not minimal, then there would be a system $\mathcal{A}'$ of dimension $m < n+1$ for f^{-1}. Applying Lemma 4.5 (Inverse Type $(1,1)$), we would get a system of dimension $m - 1 < n$ for f, contradicting minimality of $\mathcal{A}$. $\square$

Example 4.15. Taking the minimal ALS (for the anticommutator) from Example 1.19, we get, by Lemma 4.5 (Inverse Type $(1,1)$), a *minimal* ALS for $(xy+yx)^{-1}$:

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} x & -1 & . \\ y & . & -1 \\ . & y & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right).$$

Lemma 4.14 (Inverse Type $(0,0)$) gives again a minimal system for $xy + yx$.

Since it is possible to construct minimal linear representations for regular elements (see Section 3), this is in particular true for polynomials. These are of the form (4.6), since $A = I - Q$ with nilpotent Q, which can be choosen upper triangular. This can be seen either by looking at proper linear systems (Section 3) where admissible transformations are conjugations (of the system matrix such that the first component of the solution vector is left untouched) or by the following proposition.

Proposition 4.16 ([CR99, Proposition 2.1]). *Let $f \in \mathbb{K}(\!\langle X \rangle\!)$.*
   (i) f is a power series if and only if in any minimal representation, the constant term of its system matrix is invertible. There is then a minimal representation which is unital, that is, $A_0 = I$.
   (ii) f is a polynomial if and only if in any unital minimal representation, the matrix $A = A_1 \otimes x_1 + \ldots + A_d \otimes x_d$ is nilpotent. There is then a minimal representation with a unitriangular (ones on and zeros below the diagonal) system matrix.

Example 4.17. The element xyz^{-1} admits the following *minimal* ALS

$$\mathcal{A} = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -x & . \\ . & 1 & -y \\ . & . & z \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

with right family $t = (1, x, xyz^{-1})$ and left family $s = (xyz^{-1}, yz^{-1}, z^{-1})$. Now Lemma 4.8 (Inverse Type $(1,0)$) can be applied to get the minimal ALS

$$\mathcal{A}' = \left(\begin{bmatrix} 1 & . & . \end{bmatrix}, \begin{bmatrix} 1 & -z & . \\ . & y & -1 \\ . & . & x \end{bmatrix}, \begin{bmatrix} . \\ . \\ 1 \end{bmatrix} \right)$$

for $zy^{-1}x^{-1}$. Note that we can apply Lemma 4.8 again, because $1 \notin L(\mathcal{A}')$.

If an admissible linear system $\mathcal{A}$ is of the form (4.12) in Lemma 4.11 (Inverse Type $(0,1)$), then it follows immediately that $1 \in L(\mathcal{A})$. Conversely, assuming $1 \in L(\mathcal{A})$, the proof of the existence of an admissible transformation (P, Q) such that $P\mathcal{A}Q$ is of the form (4.12) is a bit more involved and requires minimality.

Lemma 4.18 (for Inverse Type $(0,1)$). *Let $\mathcal{A} = (u, A, v)$ be a minimal admissible linear system with $\dim \mathcal{A} = n \geq 2$ and $1 \in L(\mathcal{A})$. Then there exists an admissible transformation (P, Q) such that (uQ, PAQ, Pv) is of the form (4.12).*

Proof. Without loss of generality, assume that $v = [0, \ldots, 0, 1]^\top$ and the left family $s = A^{-1}v$ is $(s_1, s_2, \ldots, s_{n-1}, 1)$. Otherwise it can be brought to this form by some admissible transformation (P°, Q°). Now let $\bar{A}$ denote the upper left $(n-1) \times (n-1)$ block of A, let $\bar{s} = (s_1, \ldots, s_{n-1})$ and write $As = v$ as

$$\begin{bmatrix} \bar{A} & b \\ c & d \end{bmatrix} \begin{bmatrix} \bar{s} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

This system is equivalent to

$$\begin{bmatrix} \bar{A} & b \\ c & d-1 \end{bmatrix} \begin{bmatrix} \bar{s} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since the left family is $\mathbb{K}$-linearly independent (by minimality of $\mathcal{A}$), the matrix $\tilde{A} = \begin{bmatrix} \bar{A} & b \\ c & d-1 \end{bmatrix}$ cannot be full. We claim that there is only one possibility to transform $\tilde{A}$ to a hollow matrix, namely with zero last row. If we cannot produce a $(n-i) \times i$ block of zeros (by invertible transformations) in the first $n-1$ rows of $\tilde{A}$, then we cannot get blocks of zeros of size $(n-i+1) \times i$ and we are done.

Now assume that there are invertible matrices $P' \in \mathbb{K}^{(n-1) \times (n-1)}$ and (admissible) $Q \in \mathbb{K}^{n \times n}$ with $(Q^{-1}s)_1 = s_1$, such that $P'[\bar{A}, b]Q$ contains a zero block of size $(n-i) \times i$ for some $i = 1, \ldots, n-1$. There are two cases. If the first $n-i$ entries in column 1 cannot be made zero, we construct an upper right zero block:

$$\hat{A} = \begin{bmatrix} A_{11} & \cdot \\ A_{21} & A_{22} \end{bmatrix}, \quad \hat{s} = Q^{-1}s \quad \text{and} \quad \hat{v} = Pv = v$$

where A_{11} has size $(n-i) \times (n-i)$. If A_{11} were not full, then A would not be full (the last row is not involved in the transformation). Hence this pivot block is invertible over the free field. Therefore $\hat{s}_1 = \hat{s}_2 = \ldots = \hat{s}_{n-i} = 0$. Otherwise we construct an upper left zero block in PAQ. But then $\hat{s}_{i+1} = \hat{s}_{i+2} = \ldots = \hat{s}_n = 0$. Both contradict $\mathbb{K}$-linear independence of the left family.

So there is only one block left, which can make $\tilde{A}$ non-full. Hence, by Lemma 1.3, the modified (system) matrix $\tilde{A}$ is associated over $\mathbb{K}$ to a linear hollow matrix with a $1 \times n$ block of zeros, say in the last row (the columns and the first $n-1$ rows are left unouched):

$$\begin{bmatrix} I_{n-1} & \cdot \\ T & 1 \end{bmatrix} \begin{bmatrix} \bar{A} & b \\ c & d-1 \end{bmatrix} I_n = \begin{bmatrix} \bar{A} & b \\ 0 & 0 \end{bmatrix}.$$

Hence we have $T\bar{A} + c = 0$ and therefore

$$\begin{bmatrix} \bar{A} & b \\ 0 & Tb + d \end{bmatrix} \begin{bmatrix} \bar{s} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Clearly $Tb + d = 1$. The transformation

$$(P, Q) = \left(\begin{bmatrix} I_{n-1} & \vdots \\ T & 1 \end{bmatrix} P^\circ, Q^\circ \right)$$

does the job. $\qquad\square$

Lemma 4.19 (for Inverse Type $(1, 0)$)**.** *Let $\mathcal{A} = (u, A, v)$ be a minimal admissible linear system with $\dim \mathcal{A} = n \geq 2$ and $1 \in R(\mathcal{A})$. Then there exists an admissible transformation (P, Q) such that (uQ, PAQ, Pv) is of the form* (4.9).

Proof. The proof is similar to the previous one switching the role of left and right family. $\qquad\square$

Remark. If $1 \in R(\mathcal{A})$ for some *minimal* ALS $\mathcal{A} = (u, A, v)$, say $\dim \mathcal{A} = n$, then, by Lemma 4.19, there is an admissible transformation (P, Q) such that the first column in PAQ is $[1, 0, \ldots, 0]^\top$. So, if the first column of $A = (a_{ij})$ is not in this form, an admissible transformation can be found in two steps: Firstly, we can set up a linear system to determine an $(n-1)$-duple of scalars $(\mu_2, \mu_3, \ldots, \mu_n)$ such that $a_{i1} + \mu_2 a_{i2} + \mu_3 a_{i3} + \ldots + \mu_n a_{in}$ is in $\mathbb{K}$ for $i = 1, 2, \ldots, n$. Secondly, we use elementary row transformations (Gaussian elimination in the first column) and —if necessary— permutations to get the desired form of the first column. Together, these transformations give some (admissible) transformation (P', Q').

An analogous procedure can be applied if $1 \in L(\mathcal{A})$. And it can be combined for the case as in Lemma 4.5 (Inverse Type $(1, 1)$). It works more generally for non-minimal systems, but can fail in "pathological" cases. Compare with Example 2.5.

Theorem 4.20 (Minimal Inverse)**.** *Let $0 \neq f \in \mathbb{K}(\!(X)\!)$ be given by the minimal system $\mathcal{A} = (u, A, v)$ of dimension n. Then a minimal admissible linear system for f^{-1} is given by*

$$\mathcal{A}' = \begin{cases} (4.7) \text{ of } \dim \mathcal{A}' = n - 1 & \text{if } 1 \in L(\mathcal{A}) \text{ and } 1 \in R(\mathcal{A}), \\ (4.10) \text{ of } \dim \mathcal{A}' = n & \text{if } 1 \notin L(\mathcal{A}) \text{ and } 1 \in R(\mathcal{A}), \\ (4.13) \text{ of } \dim \mathcal{A}' = n & \text{if } 1 \in L(\mathcal{A}) \text{ and } 1 \notin R(\mathcal{A}) \text{ and} \\ (4.3) \text{ of } \dim \mathcal{A}' = n + 1 & \text{if } 1 \notin L(\mathcal{A}) \text{ and } 1 \notin R(\mathcal{A}) \end{cases}$$

provided that the necessary transformations according to Lemma 4.18 and 4.19 are done before.

Proof. See Lemma 4.5, 4.8, 4.11 and 4.14. $\qquad\square$

Acknowledgement

References

[Ami66]   S. A. Amitsur. Rational identities and applications to algebra and geometry. *J. Algebra*, 3:304–359, 1966.

[And13]   G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B):2103–2181, 2013.

[Ber78]   G. M. Bergman. The diamond lemma for ring theory. *Adv. in Math.*, 29(2):178–218, 1978.

[BMS13]   S. Belinschi, T. Mai, and R. Speicher. Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. *ArXiv e-prints*, March 2013.

[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.

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

[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.

[Coh06]   P. M. Cohn. *Free ideal rings and localization in general rings*, volume 3 of *New Mathematical Monographs.* Cambridge University Press, Cambridge, 2006.

[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.

[FR04]   M. Fortin and C. Reutenauer. Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank. *Sém. Lothar. Combin.*, 52:Art. B52f, 12 pp. (electronic), 2004.

[GGRW05] I. Gelfand, S. Gelfand, V. Retakh, and R. L. Wilson. Quasideterminants. *Adv. Math.*, 193(1):56–141, 2005.

[Hig40]   G. Higman. The units of group-rings. *Proc. London Math. Soc. (2)*, 46:231–248, 1940.

[HMV06]   J. W. Helton, S. A. McCullough, and V. Vinnikov. Noncommutative convexity arises from linear matrix inequalities. *J. Funct. Anal.*, 240(1):105–191, 2006.

[KFA69]   R. E. Kalman, P. L. Falb, and M. A. Arbib. *Topics in mathematical system theory.* McGraw-Hill Book Co., New York-Toronto, Ont.-London, 1969.

[Reu96]   C. Reutenauer. Inversion height in free fields. *Selecta Math. (N.S.)*, 2(1):93–109, 1996.

[Rob84]   M. L. Roberts. Endomorphism rings of finitely presented modules over free algebras. *J. London Math. Soc. (2)*, 30(2):197–209, 1984.

[Sch61]   M. P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

[SS78]   A. Salomaa and M. Soittola. *Automata-theoretic aspects of formal power series*. Springer-Verlag, New York-Heidelberg, 1978. Texts and Monographs in Computer Science.

[Vol15]   J. Volčič. Matrix coefficient realization theory of noncommutative rational functions. *ArXiv e-prints*, May 2015.