RAIRO-Theor. Inf. Appl. 48 (2014) 107–125
DOI: 10.1051/ita/2014004

Available online at:
www.rairo-ita.org

ON THE INVERTIBILITY OF FINITE LINEAR TRANSDUCERS

Ivone Amorim[1], António Machiavelo[1] and Rogério Reis[1]

Abstract. Linear finite transducers underlie a series of schemes for Public Key Cryptography (PKC) proposed in the 90s of the last century. The uninspiring and arid language then used, condemned these works to oblivion. Although some of these schemes were afterwards shown to be insecure, the promise of a new system of PKC relying on different complexity assumptions is still quite exciting. The algorithms there used depend heavily on the results of invertibility of linear transducers. In this paper we introduce the notion of post-initial linear transducer, which is an extension of the notion of linear finite transducer with memory, and for which the previous fundamental results on invertibility still hold. This extension enabled us to give a new method to obtain a left inverse of any invertible linear finite transducer with memory. It also plays an essencial role in the necessary and sufficient condition that we give for left invertibility of linear finite transducers.

Mathematics Subject Classification. 68Q45, 94A60.

1. Introduction

The concept of public key cryptography was introduced by Diffie, Hellman and Merkle in 1976, and, in 1978, Rivest, Shamir and Adleman presented the first public key cryptosystem, called RSA [1]. The RSA system and most of the public key cryptosystems created in the following years are based on complexity assumptions related to number theory problems, namely the factorization of integers and the discrete logarithm problem. These kinds of cryptosystems are computationally

Keywords and phrases. Linear transducers, invertibility of transducers, automata based cryptography, transducer injectivity with delay.

[1] CMUP, Faculdade de Ciências da Universidade do Porto, Portugal.
ivone.amorim@dcc.fc.up.pt; ajmachia@fc.up.pt; rvr@dcc.fc.up.pt

Article published by EDP Sciences

expensive in time as well as in space, their security relies on a very small set of problems, and improvements in algorithms to solve these problems have led to the necessity of increasing the size of the keys, which leads to higher computational costs. In a series of papers [10–13], Tao introduced a family of cryptosystems based on finite transducers, named FAPKCs. These schemes seem to be a good alternative to the classical ones, since they are computationally attractive and thus suitable for application on devices with very limited computational resources, such as satellites, cellular phones, sensor networks, and smart cards [11].

The FAPKC schemes are stream cipher schemes that can be used for encryption and signature. Roughly speaking, in these systems the private key consists of two injective transducers, namely a linear transducer and a non-linear transducer of a special kind, whose left inverses can be easily computed. Using a special product for transducers, the public key is the result of applying this operation to the original pair of transducers, thus obtaining a non-linear transducer. The crucial point is that this product is such that it is easy to obtain an inverse of the result from the inverses of the factors, while it is believed to be hard to find the inverse of the product without knowing its factorization. On the other hand, the factorization of a transducer seems, by itself, to be hard too [16].

The security of these systems relies on the difficulty of inverting nonlinear finite transducers, which is related to the difficulty of factoring matrix polynomials over $\mathbb{F}_q$, as will become apparent below. The complexity of these problems is not known, apart from the trivial fact that they are both NP-problems, exactly like the integer factoring problem that is the basis of RSA.

The invertibility theory of transducers used in the FAPKCs relies heavily on invertible linear transducers. The first study on this subject was published by Tao [7], based on the work of Massey and Slain [4] on inverses of linear sequential circuits. The notion of linear finite transducer used in that paper is a slight variant of the one Nerode studies in his seminal article [5]. While Tao calls a transducer linear if the transition and output functions are both linear on the cartesian products where they are defined, for Nerode a transducer is linear simply if the output function is linear in the first variable, for all accessible states. Also, for Nerode, a transducer has a fixed initial point, while for Tao it does not. Later on, Tao published a paper with a study on the invertibility of linear finite transducers defined over a ring [8]. Other studies were pursued some years later [2, 15–17], but as in the first paper on this subject [7], none of them presents an algorithm to invert finite linear transducers. Some authors refer to a book written in Chinese on the invertibiliy of automata, where supposedly one finds an algorithm to invert linear transducers. However, since it is not possible to have access to this algorithm, unless one understands Chinese, it is important that one has an alternative algorithm.

Moreover, the study of linear transducers and their invertibility is spread over a series of papers that sometimes do not contain proofs, or refer to papers that are not easily available to the English reader. Also, Tao has introduced the concept of transducer with memory, which is a transducer that needs some of the previous

inputs and outputs in order to transit to a new state and produce a new output. This seems to be a rather convenient class of transducers which are easy to deal with. Since they "have more information" than a transducer without memory, they are "easy" to invert. However, we have never seen, in any of the works we are aware of, the motivation for this choice. Neverthless, they play a fundamental role in FAPKCs schemes. Therefore, in this paper, we give an unified presentation of the known results, as far as we can establish, on general linear transducers, as well as on linear transducers with memory. We also simplify the language used, by introducing a more categorical point of view. As our main contribution, we introduce a new class of transducers, to which the linear transducers with memory belong, and give explicitly an algorithm to invert this kind of transducers.

This paper is organized as follows. In Section 2 we recall some basic concepts and present the most important results on the subject of invertibility of finite transducers. In Section 3 we address the problem of invertibility of linear transducers, and restate two necessary and sufficient conditions for a linear finite transducer to be invertible with a fixed delay. Section 4 is dedicated to linear finite transducers with memory. Finally, in Section 5, we introduce the notion of post-initial linear transducer (PILT), which is an extension of the notion of linear finite transducer with memory, and for which the previous fundamental results on invertibility still hold. A necessary and sufficient condition for left invertibility with fixed delay of PILTs is given. The results contained in the last Section give a way to compute inverses of invertible linear transducers with memory using the Smith Normal Form, which can be computed in deterministic polynomial time [14].

2. Preliminaries

As usual, for a finite set A, we let A^n be the set of words of length n, where $n \in \mathbb{N}_0$, and $A^0 = \{\varepsilon\}$, where ε denotes the empty word. We put $A^* = \cup_{n\geq 0} A^n$, the set of all finite words, and $A^\omega = \{a_0 a_1 \dots a_n \dots \mid a_i \in A\}$ is the set of infinite words. Finally, $|\alpha|$ denotes the length of the word α.

In what follows, *finite transducer* denotes a finite state sequential machine which, in any given state, can read a symbol from a finite set $\mathcal{X}$, the input alphabet, and in doing so produces a symbol from a finite set $\mathcal{Y}$, the output alphabet, while changing to another internal state according to certain rules. Therefore, given an initial state and an input sequence of finite length, it produces an output sequence of the same length. The formal definition of a finite state transducer is as follows.

Definition 2.1. A *finite transducer* is a quintuple $\langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$, where:

- $\mathcal{X}$ is a nonempty finite set, called the *input alphabet*;
- $\mathcal{Y}$ is a nonempty finite set, called the *output alphabet*;
- Q is a nonempty finite set called the *set of states*;
- $\delta : Q \times \mathcal{X} \to Q$, called the *state transition function*;
- $\lambda : Q \times \mathcal{X} \to \mathcal{Y}$, called the *output function.*

Let $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ be a finite transducer. The state transition function δ and the output function λ can be extended to finite words, *i.e.* elements of $\mathcal{X}^*$, recursively, as follows:

$$\begin{aligned} \delta(q, \varepsilon) &= q & \qquad \delta(q, x\alpha) &= \delta(\delta(q, x), \alpha) \\ \lambda(q, \varepsilon) &= \varepsilon & \qquad \lambda(q, x\alpha) &= \lambda(q, x)\, \lambda(\delta(q, x), \alpha), \end{aligned}$$

where $q \in Q$, $x \in \mathcal{X}$, and $\alpha \in \mathcal{X}^*$. In an analogous way, λ may be extended to $\mathcal{X}^\omega$.

From these definitions it follows that one has, for all $q \in Q, \alpha \in \mathcal{X}^*$, and for all $\beta \in \mathcal{X}^* \cup \mathcal{X}^\omega$,

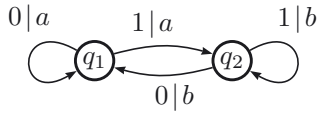
$$\lambda(q, \alpha\beta) = \lambda(q, \alpha)\, \lambda(\delta(q, \alpha), \beta).$$

A transducer can be represented by a diagram that is a digraph with labeled vertices and edges, where each state is represented by a vertex and each directed edge indicates a transition between states. The label of each edge is a compound symbol of the form i/o, where i stands for the input symbol and o for the output.

Example 1. The automaton $M = \langle \{0, 1\}, \{a, b\}, \{q_1, q_2\}, \delta, \lambda \rangle$ with

$$\begin{aligned} &\delta(q_1, 0) = q_1 \quad \delta(q_1, 1) = q_2 \quad \delta(q_2, 0) = q_1 \quad \delta(q_2, 1) = q_2 \\ &\lambda(q_1, 0) = a \quad \lambda(q_1, 1) = a \quad \lambda(q_2, 0) = b \quad \lambda(q_2, 1) = b \end{aligned}$$

is represented by the diagram:



A fundamental concept to review here is the concept of injectivity that is behind the invertibility property of the transducers used for cryptographic purposes. In fact, we will talk about two concepts: the concept of ω-injectivity and the concept of injectivity with a certain delay. These two notions of injectivity were introduced, as far as we know, by Tao, who called them weakly invertible and weakly invertible with a certain delay, respectively (see [9]). Here we use names that are more naturally related to how these terms are used in other mathematical settings.

Definition 2.2. A finite transducer $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ is *ω-injective*, if

$$\forall q \in Q,\ \forall \alpha, \alpha' \in \mathcal{X}^\omega, \quad \lambda(q, \alpha) = \lambda(q, \alpha') \implies \alpha = \alpha'.$$

That is, for any $q \in Q$, and any $\alpha \in \mathcal{X}^\omega$, α is uniquely determined by q and $\lambda(q, \alpha)$.

Definition 2.3. A finite transducer $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ is *injective with delay τ*, with $\tau \in \mathbb{N}_0$, if

$$\forall q \in Q,\ \forall x, x' \in \mathcal{X},\ \forall \alpha, \alpha' \in \mathcal{X}^\tau, \quad \lambda(q, x\alpha) = \lambda(q, x'\alpha') \implies x = x'.$$

That is, for any $q \in Q$, $x \in \mathcal{X}$, and $\alpha \in \mathcal{X}^\tau$, x is uniquely determined by q and $\lambda(q, x\alpha)$.

Below, we deal with the case $\mathcal{X} = \mathbb{F}^l$, where $\mathbb{F}$ is a field, and it will be useful to identify the elements of $\mathcal{X}^\omega$ with the elements of $\mathbb{F}[[z]]^l$, where $\mathbb{F}[[z]]$ is the ring of formal power series over $\mathbb{F}$, by replacing $x_0x_1x_2\ldots$ with $\sum_{i\geq 0} x_i z^i$. In that context, a finite transducer $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda\rangle$ is injective with delay τ if and only if

$$\lambda(q, X) \equiv \lambda(q, X') \pmod{z^{\tau+1}} \implies X \equiv X' \pmod{z}, \tag{2.1}$$

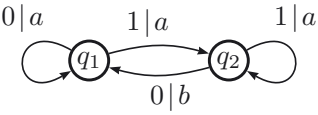
for all $X, X' \in \mathbb{F}[[z]]^l$.

Example 2. It is easy to see that for the transducer presented in Example 1 one has

$$\forall q \in \{q_1, q_2\},\ \forall x_0x_1, x_0'x_1' \in \{0,1\}^2, \quad \lambda(q, x_0x_1) = \lambda(q, x_0'x_1') \implies x_0 = x_0'.$$

Therefore, this transducer is injective with delay 1.

Example 3. The transducer induced by the diagram below is not injective with delay 1 since, for example, $\lambda(q_1, 01) = \lambda(q_1, 11)$ and $0 \neq 1$.



Theorem 2.4. *Let $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda\rangle$ be a finite transducer. If M is ω-injective, then there exists a non-negative integer $\tau \leq \frac{|Q|(|Q|-1)}{2}$ such that M is injective with delay τ.*

Proof. See Corollary 1.4.3 in [9]. □

Since every ω-injective finite transducer is injective with delay τ, for some τ on the conditions of the previous theorem, we will confine our study to these latter transducers. These are precisely the transducers used for cryptographic purposes.

Naturally, injective transducers should have inverses of some sort. In order to describe the appropriate concept, given two finite transducers $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda\rangle$ and $M' = \langle \mathcal{Y}, \mathcal{X}, Q', \delta', \lambda'\rangle$, we introduce a relation T_τ defined by

$$T_\tau = \{(q, q') \in Q \times Q' \mid \forall \alpha \in \mathcal{X}^\omega, \lambda'(q', \lambda(q, \alpha)) = \gamma\alpha \text{ for some } \gamma \in \mathcal{X}^\tau\}.$$

Hence, to say that $(q, q') \in T_\tau$ means that the state q' when fed $\lambda(q, \alpha)$ returns, after τ steps, the word α.

Remark 2.5. In this definition one may replace $\mathcal{X}^\omega$ by $\mathcal{X}^*$, but then one should also replace $\lambda'(q', \lambda(q, \alpha)) = \gamma\alpha$ by $\lambda'(q', \lambda(q, \alpha)) = \gamma\alpha'$, where α' consists of the first $|\alpha| - \tau$ characters of α.

Definition 2.6. Let $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda\rangle$ be a finite transducer. One says that M is *left invertible with delay* τ, if there is a transducer $M' = \langle \mathcal{Y}, \mathcal{X}, Q', \delta', \lambda'\rangle$ such that for all $q \in Q$, there is a $q' \in Q'$ such that $(q, q') \in T_\tau$. The transducer M' is called a *left inverse with delay* τ of M.

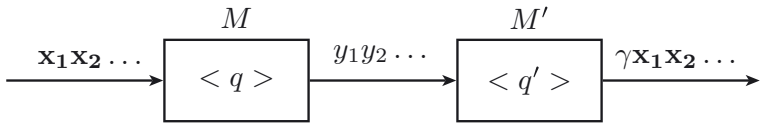
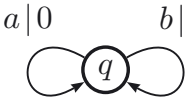


FIGURE 1. M' is a left inverse with delay τ of M when $\gamma \in \mathcal{X}^{\tau}$.

If M' is a left inverse with delay τ of M, then M' can recover the input of M with a delay of τ input symbols. Figure 1 gives a schematic representation of this concept.

Example 4. The transducer induced by the diagram below is a left inverse with delay 1 of the transducer given in Example 1.



The following result establishes the fundamental relation between the injectivity of a transducer and the existence of a left inverse. This result is presented in ([9], Cor. 1.4.4).

Theorem 2.7. *$M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ is injective with delay τ if and only if there exists a finite transducer $M' = \langle \mathcal{Y}, \mathcal{X}, Q', \delta', \lambda' \rangle$ such that M' is a left inverse with delay τ of M.*

Proof. The necessary condition is proven in ([9], Thm. 1.4.4). To prove the sufficient condition, assume that there is a transducer M' which is a left inverse of M. Let $q \in Q$, $x, x' \in \mathcal{X}$, and $\alpha, \alpha' \in \mathcal{X}^{\tau}$. Then there is a state $q' \in Q'$ such that

$$\lambda(q, x\alpha) = \lambda(q, x'\alpha') \implies \lambda'(q', \lambda(q, x\alpha)) = \lambda'(q', \lambda(q, x'\alpha')) \implies x = x'.$$

Therefore, M is injective with delay τ. □

Finally, we give the notion of a transducer with memory, in this context.

Definition 2.8. Let $\phi : \mathcal{X}^{h+1} \times \mathcal{Y}^{k} \longrightarrow \mathcal{Y}$, with $h, k \in \mathbb{N}_0$, and $\mathcal{X}, \mathcal{Y}$ two nonempty finite sets. Let $M_{\phi} = \langle \mathcal{X}, \mathcal{Y}, \mathcal{X}^{h} \times \mathcal{Y}^{k}, \delta_{\phi}, \lambda_{\phi} \rangle$ be the finite transducer given by

$$\lambda_{\phi}(\langle x_1, \ldots, x_h, y_1, \ldots, y_k \rangle, x) = \phi(x_1, \ldots, x_h, x, y_1, \ldots, y_k) = y\,;$$
$$\delta_{\phi}(\langle x_1, \ldots, x_h, y_1, \ldots, y_k \rangle, x) = \langle x_2, \ldots, x_h, x, y_2, \ldots, y_k, y \rangle,$$

for all $y_1, \ldots, y_k \in \mathcal{Y}$, $x_1, x_2, \ldots, x_h, x \in \mathcal{X}$, and where we use $< \ldots >$ to denote states of this transducer. M_{ϕ} is called the *finite transducer with memory of order (h, k) defined by ϕ.*

If, in the above definition, $(\mathcal{Y}, +)$ is a group and the function ϕ is of the form

$$\phi = f(x_1, x_2, \ldots, x_h, x_{h+1}) + g(y_1, y_2, \ldots, y_k),$$

for some $f : \mathcal{X}^{h+1} \to \mathcal{Y}$ and $g : \mathcal{Y}^k \to \mathcal{Y}$, one says that M_ϕ is a *separable finite transducer with memory*, denoted by $M_{f,g}$. Notice that, in particular, a finite transducer with no input memory is a separable finite transducer, as well as one with no output memory. The following result about separable finite transducers is mentioned in [16], without proof.

Theorem 2.9. *Let $\mathcal{Y}$ be a group, denoted additively. Then the separable transducer $M_{f,g} = \langle \mathcal{X}, \mathcal{Y}, \mathcal{X}^h \times \mathcal{Y}^k, \delta_{f,g}, \lambda_{f,g} \rangle$ is injective with delay τ if and only if the transducer $M_f = \langle \mathcal{X}, \mathcal{Y}, \mathcal{X}^h, \delta_f, \lambda_f \rangle$ is injective with delay τ.*

Proof. To simplify matters, let us identify the word $\alpha = x_1 x_2 \ldots x_n \in \mathcal{X}^n$ with the n-tuple $(x_1, x_2, \ldots, x_n)$. Then, given $q_1 \in \mathcal{X}^h$, $q_2 \in \mathcal{Y}^k$, $x \in \mathcal{X}$, one can write

$$\lambda_{f,g}(\langle q_1, q_2 \rangle, x) = f(q_1, x) + g(q_2). \tag{2.2}$$

Now, if $\alpha \in \mathcal{X}^\tau$, then $\lambda_{f,g}(\langle q_1, q_2 \rangle, x\alpha)$ is just a sequence of elements as in (2.2), and since obviously

$$f(q_1, x) + g(q_2) = f(q_1, x') + g(q_2) \iff f(q_1, x) = f(q_1, x'),$$

one concludes that

$$\lambda_{f,g}(\langle q_1, q_2 \rangle, x\alpha) = \lambda_{f,g}(\langle q_1, q_2 \rangle, x'\alpha')$$

is equivalent to

$$\lambda_f(\langle q_1 \rangle, x\alpha) = \lambda_f(\langle q_1 \rangle, x'\alpha').$$

From this the claim made follows immediately. □

This last result essentially states that the study of the injectivity with some delay of separable finite transducers can be reduced to the study of finite transducers with no output memory.

3. Linear transducers

Definition 3.1. If $\mathcal{X}, \mathcal{Y}$ and Q are vector spaces over a field $\mathbb{F}$, then a finite state transducer $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ is said to be *linear* over $\mathbb{F}$ when both $\delta : Q \times \mathcal{X} \to Q$ and $\lambda : Q \times \mathcal{X} \to Y$ are linear maps.

If $\mathcal{X}, \mathcal{Y}$, and Q have dimensions l, m and n, respectively, then there exist matrices $A \in \mathcal{M}_{n,n}(\mathbb{F})$, $B \in \mathcal{M}_{n,l}(\mathbb{F})$, $C \in \mathcal{M}_{m,n}(\mathbb{F})$, and $D \in \mathcal{M}_{m,l}(\mathbb{F})$, such that

$$\delta(q, x) = Aq + Bx,$$
$$\lambda(q, x) = Cq + Dx,$$

for all $q \in Q, x \in \mathcal{X}$. The matrices A, B, C, D are called the *structural matrices* of the finite transducer, and l, m, n are called *structural parameters* of the finite transducer.

Now, starting at a state q_0 and reading an input sequence $x_0x_1x_2\ldots$, one gets a sequence of states $q_0q_1q_2\ldots$ and a sequence of outputs $y_0y_1y_2\ldots$ satisfying the relations

$$q_{t+1} = \delta(q_t, x_t) = Aq_t + Bx_t, \tag{3.1}$$
$$y_t = \lambda(q_t, x_t) = Cq_t + Dx_t, \tag{3.2}$$

for all $t \geq 0$. Let

$$X(z) = \sum_{t\geq 0} x_t z^t, \quad Y(z) = \sum_{t\geq 0} y_t z^t, \quad Q(z) = \sum_{t\geq 0} q_t z^t,$$

regarded as elements of the $\mathbb{F}[[z]]$-modules $\mathbb{F}[[z]]^l$, $\mathbb{F}[[z]]^m$, $\mathbb{F}[[z]]^n$, respectively, where $\mathbb{F}[[z]]$ is the ring of formal power series over $\mathbb{F}$. Multiplying equality (3.1) by z^t, and adding for all $t \geq 0$, one obtains:

$$\sum_{t\geq 0} q_{t+1}z^t = AQ(z) + BX(z) \Leftrightarrow (Q(z) - q_0)z^{-1} = AQ(z) + BX(z)$$
$$\Leftrightarrow (I - Az)Q(z) = q_0 + BzX(z).$$

Since $(I - Az) \in \mathcal{M}_{n,n}(\mathbb{F})[z]$ is invertible in $\mathcal{M}_{n,n}(\mathbb{F})[[z]]$, one can rewrite the above equality as follows:

$$Q(z) = (I - Az)^{-1}q_0 + (I - Az)^{-1}BzX(z). \tag{3.3}$$

Multiplying the equality (3.2) by z^t, and adding for all $t \geq 0$, one gets:

$$Y(z) = CQ(z) + DX(z).$$

Therefore, using (3.3),

$$Y(z) = G(z)q_0 + H(z)X(z)$$

where

$$G(z) = C(I - Az)^{-1} \quad \text{and} \quad H(z) = C(I - Az)^{-1}Bz + D. \tag{3.4}$$

In [9], the matrices $G \in \mathcal{M}_{m,n}(\mathbb{F})[[z]]$ and $H \in \mathcal{M}_{m,l}(\mathbb{F})[[z]]$ are called, respectively, the *free response matrix* and the *transfer function matrix* of the transducer, designations that were suggested by [4], and that we will also use here.

The following result is presented in [15] without proof.

Theorem 3.2. *Let $M = \langle \mathbb{F}^l, \mathbb{F}^m, \mathbb{F}^n, \delta, \lambda \rangle$ be a linear finite transducer with structural matrices A, B, C and D. Let $H(z)$ be its transfer function matrix. Then, $H(z)$ is of the form*

$$\frac{1}{f(z)} \sum_{i=0}^{n} H_i z^i,$$

where $H_i \in \mathcal{M}_{m,l}(\mathbb{F})$, and $f(z) \in \mathbb{F}[z]$ is such that $f(0) = 1$.

Proof. Since

$$(I - Az)^{-1} = \frac{(I - Az)^*}{|I - Az|},$$

where P^* denotes the adjoint matrix of P, while $|P|$ denotes its determinant, one gets from (3.4) that

$$H(z) = C\frac{(I - Az)^*}{|I - Az|}Bz + D = \frac{1}{|I - Az|}\left(C(I - Az)^*Bz + |I - Az|D\right).$$

Let $f(z) = |I - Az|$, and thus $f(0) = 1$, because the independent term of $|I - Az|$ is I. Since the entries of the matrix $I - Az$ are polynomials of degree ≤ 1 and $A \in \mathcal{M}_{n,n}(\mathbb{F})$, the entries of the matrix $(I - Az)^*$ are polynomials of degree $\leq n - 1$. Also, the degree of the polynomial $|I - Az|$ is $\leq n$. Therefore, the entries of the matrix $C(I - Az)^*Bz + |I - Az|D$ are polynomials of degree $\leq n$. Since a matrix of polynomials is the same as a polynomial which the coefficients are matrices, the result follows. □

Let $\mathcal{S} = \{1 + zb(z) \mid b(z) \in \mathbb{F}[z]\}$ and $\mathbb{F}[z]_{\mathcal{S}} = \left\{\frac{f}{s} \mid f \in \mathbb{F}[z], s \in \mathcal{S}\right\}$. Then, the previous result states that the transfer function matrix of a linear finite transducer is in $\mathcal{M}(\mathbb{F}[z]_{\mathcal{S}})$. Recall that two $m \times n$ matrices A, B, with entries in a principal ideal domain R, are said to be *equivalent* if there exist an invertible matrix P in $\mathcal{M}_{m,m}(R)$ and an invertible matrix N in $\mathcal{M}_{n,n}(R)$ such that $B = PAN$. It is clear that this defines an equivalence relation in the set $\mathcal{M}_{m,n}(R)$. The following result is well known (see [3] or [6], Thm. II.9).

Theorem 3.3. *Let R be a principal ideal domain. Every matrix $A \in \mathcal{M}_{m,n}(R)$ is equivalent to a matrix of the form*

$$\mathcal{D} = \operatorname{diag}(d_1, d_2, \ldots, d_r, 0, \ldots, 0) = \begin{bmatrix} d_1 & & & & \\ & \ddots & & 0 & \\ & & d_r & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & 0 \end{bmatrix}$$

where r is the rank of A, $d_i \neq 0$ and $d_i \mid d_{i+1}$, i.e. *d_i divides d_{i+1}, for $1 \leq i \leq r-1$. The matrix $\mathcal{D}$ is called the Smith normal form of A, and the elements d_i are called the invariant factors of A.*

Since $\mathbb{F}[z]_{\mathcal{S}}$ is a principal ideal domain, and z is the unique irreducible element in $\mathbb{F}[z]_{\mathcal{S}}$, up to units, it follows from Theorem 3.3 that every matrix $H(z) \in \mathcal{M}(\mathbb{F}[z]_{\mathcal{S}})$ with rank r is equivalent to a "diagonal" matrix of the form

$$\mathcal{D}_{n_0,n_1,\ldots,n_u} = \operatorname{diag}(I_{n_0}, zI_{n_1}, \ldots, z^u I_{n_u}, 0, \ldots, 0)$$

where $n_i \geq 0$, for $0 \leq i \leq u$, $n_u \neq 0$ unless $H(z) = 0$, and $\sum_{i=0}^{u} n_i = r$. In order to facilitate the statement of the next result, we put $n_i = 0, \forall i > u$.

This is used in ([15], Thms. 1 and 2) to prove the following result, which is of great importance, because it gives two necessary and sufficient conditions for a transducer to be injective with some delay τ.

Theorem 3.4. *Let $\mathcal{X}, \mathcal{Y}$ and Q be vector spaces over a field $\mathbb{F}$, with dimensions l, m and n, respectively. Let $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda \rangle$ be a linear transducer, and let $H \in \mathcal{M}_{m,l}(\mathbb{F}[z]_{\mathcal{S}})$ be its transfer function matrix. Let $\mathcal{D} = \mathcal{D}_{n_0,n_1,\ldots,n_u}$ be the Smith normal form of H, and assume $n_u \neq 0$. Then, the following conditions are equivalent:*

i. *M is injective with delay τ;*
ii. *$\sum_{i=0}^{\tau} n_i = l$;*
iii. *there is $H' \in \mathcal{M}_{l,m}(\mathbb{F}[z]_{\mathcal{S}})$ such that $H'H = z^{\tau} I$.*

Moreover, if M is τ-injective, for some $\tau \in \mathbb{N}_0$, then it is u-injective.

Proof.

(i) $\Rightarrow$ (ii). Suppose that $\sum_{i=0}^{\tau} n_i \neq l$, that is, $\sum_{i=0}^{\tau} n_i < l$. Let $X = [0, \ldots, 0, 1]^T \in \mathcal{M}_{l,1}(\mathbb{F}[[z]])$. Then $\mathcal{D}X = \mathbf{0}$, where $\mathbf{0} = [0, \ldots, 0]^T \in \mathcal{M}_{l,1}(\mathbb{F}[[z]])$. If P and N are the invertible matrices such that $\mathcal{D} = PHN$, then $HNX = \mathbf{0}$. Putting $X' = NX$, one gets that $\lambda(0, X') = HX' = \mathbf{0}$. Since $X' \neq \mathbf{0}$, it follows that M is not injective with delay τ.

(ii) $\Rightarrow$ (iii) The hypothesis implies that, in $\mathcal{D}$, one has $\tau \geq u$ and that there are no zero columns. Now, take again P and N to be invertible matrices such that $\mathcal{D} = PHN$, and take $\mathcal{D}' = \operatorname{diag}(z^{\tau} I_{n_0}, z^{\tau-1} I_{n_1}, \ldots, z^{\tau-u} I_{n_u}) \in \mathcal{M}_{l,m}(F[z]_{\mathcal{S}})$. Then $\mathcal{D}'\mathcal{D} = z^{\tau} I$, and therefore $\mathcal{D}'PHN = z^{\tau} I$. From this it follows that $\mathcal{D}'PH = z^{\tau} N^{-1} = N^{-1} z^{\tau} I$, and hence $(N\mathcal{D}'P)H = z^{\tau} I$.

(iii) $\Rightarrow$ (i) Let q be a state of M and X, X' two input sequences such that $\lambda(q, X) \equiv \lambda(q, X') \pmod{z^{\tau+1}}$. Assume that there is $H' \in \mathcal{M}_{l,m}(F[z]_{\mathcal{S}})$ such that $H'H = z^{\tau} I$. Then, $\lambda(q, X) \equiv \lambda(q, X') \pmod{z^{\tau+1}} \iff Gq + HX \equiv Gq + HX' \pmod{z^{\tau+1}} \iff HX \equiv HX' \pmod{z^{\tau+1}} \iff H(X - X') \equiv 0 \pmod{z^{\tau+1}}$, which implies $z^{\tau} I(X - X') \equiv 0 \pmod{z^{\tau+1}}$, from which it follows $X \equiv X' \pmod{z}$. Therefore, M is injective with delay τ.

The last sentence in the statement of the theorem follows from (i) $\iff$ (ii), and the fact that $n_i = 0$, for all $i > u$. □

4. Linear transducers with memory

Let $\mathcal{X} = \mathbb{F}^l$, $\mathcal{Y} = \mathbb{F}^m$. Given $h, k \in \mathbb{N}_0$, it is easy to see that a transducer M_{ϕ} with memory of order (h, k), in the sense of Definition 2.8, is linear if and only if the function ϕ is of the form

$$\phi(x_1, x_2, \ldots, x_h, x_{h+1}, y_1, \ldots, y_k) = \sum_{i=0}^{h} a_i x_{h+1-i} + \sum_{j=1}^{k} b_j y_{k+1-j},$$

for some $a_0, \dots, a_h \in \mathcal{M}_{m,l}(\mathbb{F})$; $b_1, \dots, b_k \in \mathcal{M}_{m,m}(\mathbb{F})$, and where $x_i \in \mathcal{X}$, $y_j \in \mathcal{Y}$. Its structural matrices A, B, C, and D are as follows. Given a state q of M_ϕ, which is a vector of dimension $lh + km$ of the form

$$q = \begin{bmatrix} x_1 & \dots & x_h & y_1 & \dots & y_k \end{bmatrix}^T,$$

where $x_i \in \mathcal{M}_{1,l}(\mathbb{F})$ and $y_i \in \mathcal{M}_{1,m}(\mathbb{F})$, then putting

$$C = \begin{bmatrix} a_h & \dots & a_1 & b_k & \dots & b_1 \end{bmatrix} \text{ and } D = \begin{bmatrix} a_0 \end{bmatrix},$$

it follows that

$$\phi(x_1, \dots, x_h, x_{h+1}, y_1, \dots, y_k) = Cq + Dx_{h+1}.$$

Recalling that, by Definition 2.8,

$$\delta_\phi(\langle x_1, \dots, x_h, y_1, \dots, y_k \rangle, x) = \langle x_2, \dots, x_h, x, y_2, \dots, y_k, y \rangle,$$

where $y = \phi(x_1, \dots, x_h, x, y_1, \dots, y_k)$, if one takes

$$B = \left[\begin{array}{c:c} B_1 & B_2 \end{array}\right]^T = \left[\begin{array}{c:c} 0_{l\times(h-1)l} \;\; I_l & 0_{l\times(k-1)m} \;\; a_0^T \end{array}\right]^T,$$

and

$$A = \left[\begin{array}{c:c} A_1 & A_2 \\ \hdashline A_3 & A_4 \end{array}\right] = \left[\begin{array}{cccc:ccc} 0_l & I_l & & & & & \\ & 0_l & I_l & & & & \\ & & \ddots & \ddots & & 0_{hl\times km} & \\ & & & 0_l & I_l & & \\ 0_l & 0_l & \dots & 0_l & 0_l & & \\ \hdashline & & & & & 0_m & I_m \\ & & & & & & 0_m \;\; I_m \\ & 0_{(k-1)m\times hl} & & & & & \ddots \;\; \ddots \\ & & & & & & 0_m \;\; I_m \\ a_h & a_{h-1} & \dots & a_2 & a_1 & b_k & b_{k-1} \;\; \dots \;\; b_2 \;\; b_1 \end{array}\right]$$

it can easily be seen that $\delta_\phi(q, x) = Aq + Bx$.

Since a linear finite transducer with memory is separable, Theorem 2.9 guarantees that the injectivity of linear finite transducers with memory can be reduced to the study of linear finite transducers with no output memory.

Let now $\mathcal{L}$ be the set of all linear maps from $\mathcal{X}^{h+1}$ to $\mathcal{Y}$, for all $h \in \mathbb{N}_0$, which can be given by linear forms $\sum_{i=0}^{h} a_i x_{h-i}$ (note that $a_i \in \mathcal{M}_{m,l}(\mathbb{F})$, and $x_i \in \mathbb{F}^l$). Now, linear finite transducers with no output memory are exactly the transducers defined by functions in $\mathcal{L}$, and this set can be identified with $\mathcal{M}_{m,l}(\mathbb{F}[z]) \simeq \mathcal{M}_{m,l}(\mathbb{F})[z]$ through the map $\psi : \mathcal{M}_{m,l}(\mathbb{F}[z]) \to \mathcal{L}$ defined by

$$\psi\left(\sum_{i=0}^{h} a_i z^i\right) = \sum_{i=0}^{h} a_i x_{h-i},$$

which is clearly a bijection. Thus, in what follows, we will use indistinctly either the polynomial matrix $P = \sum_{i=0}^{h} a_i z^i$ or its linear form $\psi(P)$ to represent the linear finite transducer with no output memory defined by them. The following result states that the polynomial matrix that defines a linear finite transducer with no output memory is exactly the transfer function matrix of that transducer.

Theorem 4.1. *Let M be a linear finite transducer with memory of order $(h, 0)$, defined by $\sum_{i=0}^{h} a_i x_{h-i} \in \mathcal{L}$. Then, the transfer function matrix of M is $H = \psi^{-1}\left(\sum_{i=0}^{h} a_i x_{h-i}\right)$.*

Proof. The structural matrices of M are:

$$A = \begin{bmatrix} 0_l & I_l & & & \\ & 0_l & I_l & & \\ & & \ddots & \ddots & \\ & & & 0_l & I_l \\ & & & & 0_l \end{bmatrix}, \; B = \begin{bmatrix} 0_{(h-1)l \times l} \\ I_l \end{bmatrix},$$

$$C = \begin{bmatrix} a_h & \dots & a_1 \end{bmatrix}, \text{ and } D = \begin{bmatrix} a_0 \end{bmatrix}.$$

Then

$$I - Az = \begin{bmatrix} I_l & -zI_l & & & \\ & I_l & -zI_l & & \\ & & \ddots & \ddots & \\ & & & I_l & -zI_l \\ & & & & I_l \end{bmatrix}$$

and

$$(I - Az)^{-1} = \begin{bmatrix} I_l & zI_l & z^2 I_l & \dots & z^{h-1} I_l \\ & I_l & zI_l & & z^{h-2} I_l \\ & & \ddots & \ddots & \vdots \\ & & & I_l & zI_l \\ & & & & I_l \end{bmatrix}.$$

Consequently, the transfer function matrix of M is

$$H = C(I - Az)^{-1} Bz + D = C \begin{bmatrix} z^h I_l \\ \vdots \\ z^2 I_l \\ zI_l \end{bmatrix} + D = \sum_{i=0}^{h} a_i z^i = \psi^{-1}\left(\sum_{i=0}^{h} a_i x_{h-i}\right). \qquad \square$$

The result just proved leads to a simplification of the results of Theorem 3.4 for this kind of transducers, presented in [16] without proof, and that we now state.

Theorem 4.2. *Let M be a linear finite transducer with memory of order $(h, 0)$, defined by $E = \sum_{i=0}^{h} a_i z^i \in \mathcal{M}_{m,l}(F[z])$. Then, M is injective with some delay if and only if E has maximal rank, which when $m = l$ is, of course, equivalent to $\det(E) \neq 0$. Moreover, M is injective with delay τ if and only if the Smith normal form of E has exactly $\tau + 1$ invariant factors.*

TABLE 1. Table with the coefficients of Θ.

t	Input coefficients (ICs)					Output coefficients (OCs)				
0	a_0	$\alpha_{0,1}$	$\alpha_{0,2}$	$\dots$	$\alpha_{0,n-1}$	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	$\dots$	$\beta_{0,n}$
1	a_0	a_1	$\alpha_{1,2}$	$\dots$	$\alpha_{1,n-1}$	b_1	$\beta_{1,2}$	$\beta_{1,3}$	$\dots$	$\beta_{1,n}$
2	a_0	a_1	a_2	$\dots$	$\alpha_{2,n-1}$	b_1	b_2	$\beta_{2,3}$	$\dots$	$\beta_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$n-1$	a_0	a_1	a_2	$\dots$	a_{n-1}	b_1	b_2	b_3	$\dots$	$\beta_{n-1,n}$
$\geq n$	a_0	a_1	a_2	$\dots$	a_{n-1}	b_1	b_2	b_3	$\dots$	b_n

Proof. From the previous result, the matrix E is the transfer function matrix of M. Let $\mathcal{D} = \mathcal{D}_{n_0,n_1,\dots,n_u}$ be the Smith normal form of E, with $n_u \neq 0$. Then, from the last statement of Theorem 3.4 one concludes that M is injective with some delay if M is u-injective. Then, from the equivalence (i) $\iff$ (ii) one can conlude that M is injective with some delay if and only if $\sum_{i=0}^{u} n_i = l$, that is, E has maximal rank. □

5. Post-initial linear transducers

The search for inverses of linear transducers led us to a new class of transducers that includes the linear transducers with memory as given in Definition 2.8, and for which all the previous results still apply. We call them PILTs (post-initial linear transducers), and in this Section we show, among other things, that a left invertible PILT has a left inverse that is also a PILT.

Let $\mathbb{F}$ be a finite field, $\mathcal{X} \simeq \mathbb{F}^l$ and $\mathcal{Y} \simeq \mathbb{F}^m$, and $V = \mathcal{M}_{m,l}(\mathbb{F})$, $R = \mathcal{M}_{m,m}(\mathbb{F})$. In what follows we will regard $\mathcal{Y}$ and V as left R-modules. Consider the map $\Theta : \mathcal{X}^\omega \to \mathcal{Y}^\omega$ given by

$$y_t = \sum_{i=1}^{n} (\alpha_{t,i-1}\, x_{t+1-i} + \beta_{t,i}\, y_{t-i}) \qquad (t \geq 0) \tag{5.1}$$

where, $n \in \mathbb{N}$, $\alpha_{t,i-1} \in V, \beta_{t,i} \in R$, and

$$\forall t \geq i-1, \quad \alpha_{t,i-1} = a_{i-1} \quad \text{and} \quad \forall t \geq i, \quad \beta_{t,i} = b_i,$$

with $a_{i-1} \in V$, $b_i \in R$, for $i \in \{1, \dots, n\}$. The variables with negative indices are free and the map Θ is determined by their values, which one should think of as a set of *initial values*. The map Θ is determined by the array of constants (its coefficients) presented in Table 1, together with those initial values.

For any given set of initial values, the corresponding map Θ is a linear affine map of vector spaces over $\mathbb{F}$, and in the case they are all zero it is, of course, linear, and the fact that the sequences $(\alpha_{t,i})_t$ and $(\beta_{t,i})_t$ are eventually constant implies that Θ is then what Nerode calls an *automaton transformation*, *i.e.* is induced by a finite transducer, by a straightforward generalization of ([5], Lem. 3) to our setting. We note that this result still holds in the general case of arbitrary initial values,

TABLE 2. Coefficients of the PILT given in Example 5.

t	ICs		OCs	
0	2	1	2	0
≥ 1	1	1	1	1

since one can still use the same argument as in ([5], Lem. 3) to show that Θ has a finite number of what Nerode calls *intrinsic states*, and then ([5], Lem. 2) applies. These initial values that can also be thought of as states of the transducer, using a construction completely analogous to the transducer with memory of Tao [9].

All of the above shows that the following definition makes sense.

Definition 5.1. A post-initial linear transducer (PILT) is a transducer induced by a recurrence relation as in (5.1). If h is the largest value of $i \in \{1, \ldots, n\}$ such that $\alpha_{t,i-1} \neq 0, \forall t \leq i-1$, and k is the largest value of $j \in \{1, \ldots, n\}$ such that $\beta_{t,j} \neq 0, \forall t \leq j$, then one calls the corresponding transducer a PILT with memory (h, k).

Observation: If one represents a PILT with order (h, k) by a table similar to Table 1, then h is the index minus 1 of the highest column containing the input coefficients that has a non-zero entry. And, k is the index of the highest column containing the otput coefficients that has a non-zero entry. Of course, the linear finite transducers with memory defined in the previous Section correspond to the special case where the sequences $(\alpha_{t,i})_t$ and $(\beta_{t,i})_t$ are constant.

Example 5. Taking $n = 2$, $\mathcal{X} = \mathcal{Y} = \mathbb{F}_3$, and the map given by the coefficients on Table 2, one gets a PILT induced by the following recurrence relation:

$$\begin{cases} y_0 = 2x_0 + x_{-1} + 2y_{-1} \\ y_t = x_t + x_{t-1} + y_{t-1} + y_{t-2}, \ t \geq 1, \end{cases}$$

which has memory of order $(1, 2)$. Taking, for example, $q = \langle 1, 2, 0 \rangle$, one has $\lambda(q, 11201) = 02200$.

In what follows, given a set S, we use the notation $\mathcal{P}_n(S[z])$ to denote the set of polynomial with coefficients in S that have degree less than n.

Put $X(z) = \sum_{t \geq 0} x_t z^t \in \mathbb{F}^l[[z]] \simeq \mathbb{F}[[z]]^l$ and $Y(z) = \sum_{t \geq 0} y_t z^t \in \mathbb{F}^m[[z]] \simeq \mathbb{F}[[z]]^m$. Multiplying (5.1) by z^t and adding for all $t \geq 0$, one obtains

$$g(z)Y(z) - f(z)X(z) = r(q), \tag{5.2}$$

where $g(z) = I - \sum_{i=1}^{n} b_i z^i \in \mathcal{P}_{n+1}(R[z])$, $f(z) = \sum_{i=0}^{n} a_i \, z^i \in \mathcal{P}_{n+1}(V[z])$, and $r : Q \to \mathcal{P}_n(\mathbb{F}[z]^m)$ is given by:

$$r(q) = \sum_{t=0}^{n-1} \left(\sum_{i=t+2}^{n} \alpha_{t,i-1} x_{t+1-i} + \sum_{i=t+1}^{n} \beta_{t,i} y_{t-i} \right) z^t, \tag{5.3}$$

TABLE 3. Coefficients of the PILT given in Example 6.

t	ICs			OCs		
0	$\begin{bmatrix}0&1\\0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}0&0\\0&0\\0&0\end{bmatrix}$	$\begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$
1	$\begin{bmatrix}0&1\\0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}1&1\\1&1\\1&1\end{bmatrix}$	$\begin{bmatrix}1&1\\0&0\\0&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&0&0\\1&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$
≥ 2	$\begin{bmatrix}0&1\\0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}1&1\\1&1\\1&1\end{bmatrix}$	$\begin{bmatrix}1&0\\0&0\\0&0\end{bmatrix}$	$\begin{bmatrix}1&0&0\\0&0&0\\1&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$	$\begin{bmatrix}1&1&0\\0&0&0\\0&1&0\end{bmatrix}$

if $q = \langle x_{-(n-1)}, \ldots, x_{-1}, y_{-n}, \ldots, y_{-1}\rangle$. We will say that q gives *the initial conditions*, or the *initial state*.

It is clear that the two forms of inducing a transducer, either by an equation of the form (5.1) or by one of the form (5.2), are equivalent.

Example 6. Take $l = 2$, $m = 3$ and $n = 3$. Let $M = \langle (\mathbb{F}_2)^2, (\mathbb{F}_2)^3, (\mathbb{F}_2)^5, \delta, \lambda, \rangle$ be the PILT with memory of order $(2, 3)$ induced by

$$y_t = \sum_{i=1}^{3} (\alpha_{t,i-1}\, x_{t+1-i} + \beta_{t,i}\, y_{t-i}) \qquad (t \geq 0) \tag{5.4}$$

whose coefficients are given in Table 3.
Then, the series of inputs and outputs that satisfy this equation are the same as the ones that satisfy

$$g(z)Y(z) - f(z)X(z) = r(q),$$

with

$$f(z) = \begin{bmatrix}0&1\\0&0\\0&1\end{bmatrix} + \begin{bmatrix}1&1\\1&1\\1&1\end{bmatrix} z + \begin{bmatrix}1&0\\0&0\\0&0\end{bmatrix} z^2,$$

$$g(z) = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} + \begin{bmatrix}1&0&0\\0&0&0\\1&0&0\end{bmatrix} z + \begin{bmatrix}1&1&0\\0&0&0\\0&1&0\end{bmatrix} z^3,$$

$$r(q) = \begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix} x_{-2} + \begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix} y_{-3} + \begin{bmatrix}1&1\\0&0\\0&0\end{bmatrix} x_{-1} z + \begin{bmatrix}1&1&0\\0&0&0\\0&1&0\end{bmatrix} y_{-1} z^2.$$

We are now ready to state a result that will allow us to give a necessary and sufficient condition for the left invertibility of PILTs, and consequently of linear transducers with memory.

Proposition 5.2. *Let $f \in \mathcal{M}_{m,l}(\mathbb{F})[z]$, $g \in \mathcal{M}_{m,m}(\mathbb{F})[z]$ with $g(0) = I$, and let $r : Q \to \mathbb{F}[z]^m$ be given by an expression of the form (5.3), and let $M = \langle \mathcal{X}, \mathcal{Y}, Q, \delta, \lambda\rangle$ be a PILT induced by the equation $gY - fX = r(q)$, as described above. Then the*

series of inputs and outputs of M, for some initial conditions q, satisfy an equation of the form $uX - vY = s$, for some $u \in \mathcal{M}_{l,l}(\mathbb{F})[z]$ with $u \equiv z^\tau I \pmod{z^{\tau+1}}$, $v \in \mathcal{M}_{l,m}(\mathbb{F})[z]$, and $s \in \mathbb{F}[z]^l$, if and only if

$$\exists\, p \in \mathcal{M}_{l,m}(\mathbb{F})[z] :\ pf \equiv z^\tau I \pmod{z^{\tau+1}}.$$

Proof. One direction is obvious. If there exists $p \in \mathcal{M}_{l,m}(\mathbb{F})[z]$ such that $pf \equiv z^\tau I \pmod{z^{\tau+1}}$, then just by multiplying both sides of equation $gY - fX = r(q)$ by p, on the left, one immediately gets the desired result.

To prove the only if part, assume that there are u, v, s in the conditions described in the statement of the theorem. Since $u \equiv z^\tau I \pmod{z^{\tau+1}}$, there is a polynomial w, such that $u = z^\tau w$ and $w(0) = I$.

From $gY - fX = r(q)$ and $g(0) = I$, one gets $Y = g^{-1}fX + g^{-1}r(q)$. Substituting this into $uX - vY = s$, one gets

$$\left(z^\tau w - vg^{-1}f\right) X = vg^{-1}r(q) + s.$$

Since this must be true for all $X \in \mathcal{X}^\omega \simeq \mathbb{F}[[z]]^l$, it follows that $z^\tau w - vg^{-1}f$ must be the zero matrix, which then implies that

$$z^\tau I = w^{-1}vg^{-1}f,$$

where I is the identity matrix of the appropriate size. Moreover, since f is a "polynomial", one concludes that $w^{-1}vg^{-1}$ is also a "polynomial", more precisely, an element of $\mathcal{M}_{l,m}(\mathbb{F})[z]$. Therefore, putting $p = w^{-1}vg^{-1}$, one gets the claimed result. □

Theorem 5.3. *Let M be a PILT induced by $f \in \mathcal{M}_{m,l}(\mathbb{F})[z]$, $g \in \mathcal{M}_{m,m}(\mathbb{F})[z]$ with $g(0) = I$, and $r : Q \to \mathbb{F}[z]^m$, as above. Then M has a left inverse with delay τ if and only if*

$$\exists\, p \in \mathcal{M}_{l,m}(\mathbb{F})[z] :\ pf \equiv z^\tau I \pmod{z^{\tau+1}}.$$

In that case, if $w \in \mathcal{M}_{l,l}(\mathbb{F})[z]$ is such that $pf = z^\tau w$, with $w(0) = I$, then an inverse with delay τ of M is the transducer induced by $wY - pgX = -pr(q)$.

Proof. Suppose M has a left inverse with delay τ, M'. Let $wY - vX = s(q')$, with $w(0) = I$, be an equation that induces M'. Then, for any input-output pair (X, Y) of M, and for any initial conditions q, there are initial conditions q' of M' and a polynomial $\gamma \in \mathcal{P}_\tau(\mathbb{F}[z])^l$ such that $(Y, z^\tau X + \gamma)$ is an input-output pair of M'. This implies that

$$wz^\tau X - vY = s(q') - w\gamma,$$

and the previous proposition then applies to show the desired congruence.

Conversely, assume the existence of p as stated, and let u be such that $pf = z^\tau u$. Then $u(0) = I$, and multiplying by p the equation defining M, one gets:

$$pgY - pfX = pr(q), \tag{5.5}$$

which is equivalent to

$$u\,(z^\tau X) - pgY = -pr(q), \tag{5.6}$$

where $-pr(q)$ is an expression of the form

$$\sum_{t=0}^{n-1} \varphi_t(x_{-(n-1)}, \ldots, x_{-1}, y_{-n}, \ldots, y_{-1}) z^t,$$

with $\varphi_t(x_{-(n-1)}, \ldots, x_{-1}, y_{-n}, \ldots, y_{-1})$ being linear forms. Now, any expression of this form can be written in the form (5.3) by introducing new variables with zero coefficients, if necessary. This is better seen through an example – see Example 7 below. It follows that $-pr(q) = s(q')$ for some $s : Q' \to \mathbb{F}[z]^l$ of form (5.3) and $q' \in Q$. Since equation (5.5) is verified for any input-output pair (X, Y) of M, one concludes that the transducer M' defined by $uY - pgX = s(q')$ is a left inverse of M with delay τ. □

Note that the left inverse whose existence is here shown outputs zeros before starting to recover the input. Also, this result immediately gives an algorithm to find such inverse, namely:

(1) find $p \in \mathcal{M}_{\ell,m}(\mathbb{F})[z]$ such that $pf \equiv z^\tau I \pmod{z^{\tau+1}}$;
(2) determine $w \in \mathcal{M}_{\ell,\ell}(\mathbb{F})[z]$ such that $pf = z^\tau w$, with $w(0) = I$;
(3) compute pg and $pr(q)$;

To find p satisfying (5) one uses the proof of the following result, in which $\mathcal{M}(S)$ will denote the union of all rings of matrices over the ring S.

Theorem 5.4. *Let $\mathbb{F}$ be a field, and $F \in \mathcal{M}(\mathbb{F}[z])$. Then*

$$\exists\, P \in \mathcal{M}(\mathbb{F}[z]) :\ PF \equiv z^\tau I \pmod{z^{\tau+1}} \iff z^{\tau+1} \nmid d,$$

where d is the invariant factor with the highest degree of F in Smith's normal form, and I is the appropriate identity matrix.

Proof. Let $F \in \mathcal{M}(\mathbb{F}[z])$. Since $\mathbb{F}[z]$ is a principal ideal domain, there exist invertible matrices $U, V \in \mathcal{M}(\mathbb{F}[z])$, with the appropriate dimensions, and such that $D = UFV$ is the Smith's normal form of F. One then has,

$$\begin{aligned}
&\exists\, P \in \mathcal{M}(\mathbb{F}[z]) :\ PF \equiv z^\tau I \pmod{z^{\tau+1}} \Leftrightarrow \\
\Leftrightarrow\ &\exists\, P \in \mathcal{M}(\mathbb{F}[z]) :\ PU^{-1}UFV \equiv z^\tau V \pmod{z^{\tau+1}} \Leftrightarrow \\
\Leftrightarrow\ &\exists\, P \in \mathcal{M}(\mathbb{F}[z]) :\ V^{-1}PU^{-1}D \equiv z^\tau I \pmod{z^{\tau+1}} \Leftrightarrow \\
\Leftrightarrow\ &\exists\, P = (h_{ij})_{i,j} \in \mathcal{M}(\mathbb{F}[z]) :\ PD \equiv z^\tau I \pmod{z^{\tau+1}} \Leftrightarrow \\
\Leftrightarrow\ &\forall_{i,j}\, \exists\, p_{i,j} \in \mathbb{F}[z] :\ \begin{cases} p_{ij} \equiv 0 \pmod{z^{\tau+1}}, & \text{if } i \neq j \\ p_{ii} d_i \equiv z^\tau \pmod{z^{\tau+1}}, & \text{otherwise.} \end{cases} \Leftrightarrow \\
\Leftrightarrow\ &z^{\tau+1} \nmid d,
\end{aligned}$$

where d_i are the invariant factors of F, and d is the one with the highest degree. □

This yields the following algorithm to find p satisfying (5) in the previous algorithm.

(1) Determine the matrices U, V such that $UFV = \mathrm{SNF}(F)$.

(2) Construct a matrix A such that $\begin{cases} a_{ij} \equiv 0 \pmod{z^{\tau+1}}, & \text{if } i \neq j \\ a_{ii} d_i \equiv z^{\tau} \pmod{z^{\tau+1}}, & \text{otherwise.} \end{cases}$

(3) Compute $P = VAU$.

Example 7 (Continuing Example 6)**.** Let F be the polynomial matrix that corresponds to $f(z)$. The Smith normal form of F and the matrices U, V such that $D = UFV$ are:

$$D = \begin{bmatrix} 1 & 0 \\ 0 & z \\ 0 & 0 \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & 1 & 0 \\ 0 & z+1 & z \\ 1 & z^2+z & z^2+1 \end{bmatrix}, \qquad V = \begin{bmatrix} 0 & 1 \\ 1 & z^2 \end{bmatrix}.$$

Choosing $\tau = 1$, and since $z^2 \nmid z$, by the previous theorem, there exists $P \in \mathcal{M}(\mathbb{F}[z])$ such that $PF \equiv z^{\tau} I \pmod{z^{\tau+1}}$, and using the algorithm just mentioned, one has, for example:

$$P = \begin{bmatrix} 0 & z+1 & z \\ z & z & 0 \end{bmatrix}.$$

Let $p(z)$ be the matrix polynomial that corresponds to P. Taking $u = pf$, $v = pg$ and $s(q') = -pr(q)$, one gets that the series of inputs and outputs of M satisfy the equation $uX - vY = s(q')$, with

$$\begin{aligned} u &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} z^3, \\ v &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} z + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} z^2 + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} z^4, \\ s(q') &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} y_{-3} + \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_{-2} + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} y_{-3} \right) z + \\ &\quad + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} x_{-1} z^2 + \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} y_{-1} z^3. \end{aligned}$$

Therefore, by the same result, a left inverse with delay 1 of M is the post-initial linear transducer $M' = \left\langle (\mathbb{F}_2)^3, (\mathbb{F}_2)^2, (\mathbb{F}_2)^6, \delta', \lambda' \right\rangle$ with memory of order $(4, 2)$ induced by the equation $wY - vX = s(q')$, where $v, s(q')$ are as defined above, and

$$w = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} z^2.$$

Finally, from Proposition 5.2, Theorems 5.3 and 5.4 one gets the following necessary and sufficient condition for the left invertibility of PILTs.

Corollary 5.5. *Let $\mathbb{F}$ be a field. Let $f \in \mathcal{M}_{m,l}(\mathbb{F})[z]$, $g \in \mathcal{M}_{m,m}(\mathbb{F})[z]$ such that $g(0) = I$, and $r : Q \to \mathbb{F}[z]^m$ given by an expression of the form (5.3). Let $M = \left\langle \mathbb{F}^l, \mathbb{F}^m, Q, \delta, \lambda \right\rangle$ be a PILT induced by the equation $gY - fX = r(q)$, as described above. Then, M is left invertible with delay τ if and only if*

$$z^{\tau+1} \nmid d,$$

where d is the invariant factor with the highest degree of f, when f is seen as an element of $\mathcal{M}_{m,l}(\mathbb{F}[z])$.

6. Conclusion

In this paper we give an account of some of the known results on the invertibility of linear finite transducers. By considering an appropriate extension of the notion of linear transducer with memory, and working on rings of formal power series and some associated modules, we were able to get a necessary and sufficient condition for the invertibility of linear transducers with memory. We also gave a way to compute inverses of any invertible linear transducer with memory, using the Smith Normal Form for polynomial matrices.

Acknowledgements. This work was partially funded by the European Regional Development Fund through the programme COMPETE and by the Portuguese Government through the FCT under the Projects PEst-C/MAT/UI0144/2011 and CANTE-PTDC/EIA-CCO/101904/2008. Ivone Amorim is funded by FCT Grant SFRH/BD/84901/2012. The authors gratefully acknowledge the useful suggestions and comments of the unknown referees.

References

[1] W. Diffie, The First Ten Years of Public-Key Cryptography. *Proc. IEEE* **76** (1988) 560–577.
[2] O. Haiwen and D. Zongduo, Self-Injective Rings and Linear (Weak) Inverses of Linear Finite Automata over Rings. *Science in China, Series A* **42** (1999) 140–146.
[3] N. Jacobson, *Basic Algebra I.* W H Freeman & Co (1985).
[4] J.L. Massey and M.K. Slain, Inverses of Linear Sequential Circuits. *IEEE Trans. Comput.* **C-17** (1968) 330–337.
[5] A. Nerode, Linear Automaton Transformations. *Proc. Amer. Math. Soc.* **9** (1958) 541–544.
[6] M. Newman, *Integral Matrices.* Academic Press (1972).
[7] R. Tao, Invertible Linear Finite Automata. *Sci. Sinica* **XVI** (1973) 565–581.
[8] R. Tao, Invertibility of Linear Finite Automata Over a Ring. *Automata, Languages and Programming*, in vol. 317 of *Lect. Notes Comput. Sci.* Springer Berlin, Heidelberg (1988) 489–501.
[9] R. Tao, *Finite Automata and Application to Cryptography.* Springer Publishing Company, Incorporated (2009).
[10] R. Tao and S. Chen, A Finite Automaton Public Key Cryptosystem and Digital Signatures. *Chinese J. Comput.* **8** (1985) 401–409. (in Chinese).
[11] R. Tao and S. Chen, A Variant of the Public Key Cryptosystem FAPKC3. *J. Netw. Comput. Appl.* 20 (1997) 283–303.
[12] R. Tao and S. Chen, The Generalization of Public Key Cryptosystem FAPKC4. *Chinese Sci. Bull.* **44** (1999) 784–790.
[13] R. Tao, S. Chen and C. Xuemei, FAPKC3: A New Finite Automaton Public Key Cryptosystem. *J. Comput. Sci. Techn.* **12** (1997) 289–305.
[14] G. Villard, Generalized subresultants for computing the Smith normal form of polynomial matrices. *J. Symb. Comput.* **20** (1995) 269–286.
[15] D. Zongduo and Y. Dingfengd, Weak Invertibility of Linear Finite Automata I, Classification and Enumeration of Transfer Functions. *Sci. In China (Series A)* **39** (1996) 613–623.
[16] D. Zongduo, Y. Dingfeng and K.Y. Lam, Weak Invertibility of Finite Automata and Cryptanalysis on FAPKC. *Advances in Cryptology – AsiaCrypt'98*, in vol. 1514 of *Lect. Notes Comput. Sci.* Edited by K. Ohta and D. Pei. Springer-Verlag (1998) 227–241.
[17] D. Zongduo, Y. Dingfengd, Z. Qibin and O. Haiwen, Classification and Enumeration of Matched Free Response Matrices of Linear Finite Automata. *Acta Math. Sinica, New Ser.* **13** (1997) 133–144.

Communicated by M. Holzer.
Received January 31, 2013. Accepted January 31, 2014.