Duality between Multidimensional Convolutional Codes and Systems*

Heide Gluesing-Luerssen
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556-5683, USA
e-mail: gluesing-luerssen.1@nd.edu
or gluesing@mathematik.uni-oldenburg.de

Joachim Rosenthal
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556-5683, USA
e-mail: Rosenthal.1@nd.edu

Paul A. Weiner
Department of Mathematics
Saint Mary's University of Minnesota
Winona, MN 55987, USA
e-mail: pweiner@smumn.edu

June 2000

Abstract: Multidimensional convolutional codes arise as a generalization of "classical" onedimensional codes. We will introduce m-dimensional convolutional codes of length n as submodules of $\mathcal{D}^n$ where $\mathcal{D}$ is the polynomial ring in m variables over a finite field. Besides their coding theoretic significance, they can also be regarded as the annihilating modules of systems of partial difference equations, the latter being studied in much detail in discrete-time multidimensional systems theory. We will apply the duality theorem of Oberst [5] to this particular case and employ the duality to investigate certain first-order representations of onedimensional convolutional codes.

Dedicated to Diederich Hinrichsen on the occasion of his 60th birthday

*Supported in part by NSF grant DMS-96-10389. Paul Weiner would like to thank the Center for Applied Mathematics at Notre Dame for a fellowship which financially supported the presented research.

1 Introduction

Data transmission over noisy channels requires implementation of good coding devices. Convolutional codes belong to the most widely implemented codes. These codes represent in essence discrete time linear systems over a fixed finite field $\mathbb{F}$. Because of this reason a study of convolutional codes requires a good understanding of techniques from linear systems theory.

Multidimensional convolutional codes generalize (one dimensional) convolutional codes and they correspond to multidimensional systems widely studied in the systems literature. (See [5] and its references). These codes are very suitable if e.g. the data transmission requires the encoding of a sequence of pictures and we will explain this at the end of this section.

In the sequel we will assume that a certain message source is already encoded through a sequence of vectors $m_i \in \mathbb{F}^k$, $i = 1, \dots, \gamma$. If every vector in $\mathbb{F}^k$ is a valid message word, then the change of one coordinate of a vector $m \in \mathbb{F}^k$ will result in another valid message vector $\tilde{m} \in \mathbb{F}^k$ and the error can neither be detected nor corrected. In order to overcome this difficulty one can add some redundancy by constructing an injective linear map

$$\varphi : \mathbb{F}^k \longrightarrow \mathbb{F}^n$$

having the property that the Hamming distance $\operatorname{dist}(\varphi(m_1), \varphi(m_2))$, that is the number of different entries in the vectors $\varphi(m_1)$ and $\varphi(m_2)$, is at least d whenever $m_1 \neq m_2$. If one transmits the n-vector $\varphi(m_1)$ instead of the k-vector m_1 then it is possible to correct up to $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors for every transmitted n-vector, for details see Lemma 2.8.

The linear transformation φ defines an encoder and $\operatorname{im}(\varphi) \subset \mathbb{F}^n$ is called a linear block code. In order to describe the encoding of a whole sequence of message words $m_0, m_1, \dots, m_\gamma \in \mathbb{F}^k$ it will be convenient to introduce the polynomial vector $m(z) := \sum_{i=1}^{\gamma} m_i z^i \in \mathbb{F}^k[z]$. The encoding procedure is then compactly written by:

$$\hat{\varphi} : \mathbb{F}^k[z] \longrightarrow \mathbb{F}^n[z], \quad m(z) \longmapsto \hat{\varphi}(m(z)) = \sum_{i=1}^{\gamma} \varphi(m_i) z^i.$$

If $\mathcal{D}$ denotes the polynomial ring $\mathcal{D} = \mathbb{F}[z]$ then one immediately verifies that $\hat{\varphi}$ describes an injective module homomorphism between the free modules $\mathcal{D}^k$ and $\mathcal{D}^n$ and $\operatorname{im}\left(\hat{\varphi}(\mathcal{D}^k)\right) \subset \mathcal{D}^n$ is a submodule.

In general not every injective module homomorphism between $\mathcal{D}^k$ and $\mathcal{D}^n$ is of this form. Indeed $\hat{\varphi}$ has the peculiar property that the i-th term of $\hat{\varphi}(m(z))$ only depends on the i-th term of $m(z)$. In other words the encoder $\hat{\varphi}$ has 'no memory'. In general it is highly desirable to invoke encoding schemes where $\hat{\varphi} : \mathcal{D}^k \to \mathcal{D}^n$ is an arbitrary injective module homomorphism. The image of such a module homomorphism is then called a (1-D) convolutional code.

1-D convolutional codes are very much suited in the encoding of sequences of message blocks. Sometimes it might be desirable that the data is represented through polynomial rings in several variables. This leads us then to the definition of a m-dimensional convolutional code whose basic properties we intend to study in this paper.

The following example will illustrate the usefulness of multidimensional convolutional codes.

Example 1.1 Let $\mathcal{D} = \mathbb{F}[z_1, z_2, z_3]$ be the polynomial ring in the indeterminates z_1, z_2, z_3. A whole motion picture (without sound) can be described by one element of $\mathcal{D}^k$. Indeed if $f \in \mathcal{D}^k$,

$$f(z_1, z_2, z_3) = \sum_{x=0}^{\xi} \sum_{y=0}^{\rho} \sum_{t=0}^{\tau} f_{(x,y,t)} z_1^x z_2^y z_3^t \in \mathbb{F}^k[z_1, z_2, z_3]$$

then we can view the vector $f_{(x,y,t)} \in \mathbb{F}^k$ as describing the color and the intensity of a pixel point with coordinates (x, y) at time t.

In practice the encoding of the element $f(z_1, z_2, z_3) \in \mathcal{D}^k$ is done in the following way. At a particular time instance t all the data vectors $f_{(x,y,t)}$ are combined into a large vector $\hat{f}_t \in \mathbb{F}^K$, where K depends on the size of k and the number of pixel points on the screen. In this way we can identify each element $f(z_1, z_2, z_3) \in \mathcal{D}^k$ of above type with a polynomial vector $\hat{f}(z_3) \in \mathbb{F}^K[z_3]$. The vector $\hat{f}(z_3)$ is then encoded with a usual 1-D encoding scheme. This encoding scheme is shift invariant with respect to time but it is in general not shift invariant with respect to the z_1 and z_2 directions on the screen.

In order to achieve an encoding scheme which is also shift-invariant with respect to the coordinate axes of the screen one can do the following: Construct an injective module homomorphism $\hat{\varphi} : \mathcal{D}^k \to \mathcal{D}^n$. The image then describes a 3-dimensional convolutional code which is invariant with respect to time and both coordinate axes. The transmission of an element $\mathcal{D}^n$ is then done by choosing a term order among the monomials of the form $z_1^x z_2^y z_3^t$.

2 Multidimensional Convolutional Codes

In this section we introduce multidimensional convolutional codes as submodules of $\mathcal{D}^n$, where $\mathcal{D}$ denotes a polynomial ring in m variables. Our presentation in this section follows closely [12, Chapter 2].

We begin by setting some notations. Let $\mathbb{F}$ be any finite field and define $\mathcal{D} = \mathbb{F}[z_1, \ldots, z_m]$ to be the polynomial ring in m indeterminates over $\mathbb{F}$. We will mainly use the shorter form

$$\mathcal{D} = \mathbb{F}[z] = \Big\{ \sum_{\alpha \in \mathbb{N}^m}{}' f_\alpha z^\alpha \,\Big|\, f_\alpha \in \mathbb{F} \Big\},$$

where for $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$ the notation z^α stands for $z_1^{\alpha_1} \cdot \ldots \cdot z_m^{\alpha_m}$ and where $\sum'$ means this sum being finite. Note that $\mathcal{D}$ is $\mathbb{F}$-isomorphic to the m-dimensional finite sequence space

$$\mathcal{S} = \{ f : \mathbb{N}^m \longrightarrow \mathbb{F} \mid f \text{ has finite support} \},$$

the isomorphism given by

$$\begin{aligned} \psi \,:\, \mathcal{S} &\longrightarrow \mathcal{D} \\ f &\longmapsto \sum_{\alpha \in \mathbb{N}^m} f(\alpha) z^\alpha \end{aligned}$$

One can visualize the elements of $\mathcal{S}$ by using the integer lattice of the first quadrant of $\mathbb{R}^m$ and attaching the element $f(\alpha_1, \ldots, \alpha_m) \in \mathbb{F}$ to the point with coordinates $(\alpha_1, \ldots, \alpha_m)$. It is convenient to omit the attachment if $f(\alpha_1, \ldots, \alpha_m) = 0$.

Example 2.1 We visualize the polynomial $f(z_1, z_2) = 1 + 2z_1^2 + 2z_1 z_2 \in \mathbb{F}_3[z_1, z_2]$ as well as $z_1 f(z_1, z_2)$.

$$\begin{array}{cccccc}
\bullet & \bullet_2 & \bullet & \bullet & \xrightarrow{\psi} & 1 + 2{z_1}^2 + 2z_1 z_2 \\
\bullet_1 & \bullet & \bullet_2 & \bullet & & \\
 & & & & & \downarrow z_1\cdot \\
\bullet & \bullet & \bullet_2 & \bullet & \xleftarrow{\psi^{-1}} & z_1 + 2{z_1}^3 + 2{z_1}^2 z_2\ . \\
\bullet & \bullet_1 & \bullet & \bullet_2 & &
\end{array}$$

As the example indicates, multiplication with z_i in the ring $\mathcal{D}$ corresponds to the forward shift along the ith axis in $\mathcal{S}$. This can be verified with the help of the following commutative diagram

$$\begin{tikzcd}
\big(f(\alpha)\big)_{\alpha \in \mathbb{N}^m} \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \sum_{\alpha \in \mathbb{N}^m} f(\alpha) z^\alpha \arrow[ddd, mapsto] \\
& \mathcal{S} \arrow[r, "\psi"] \arrow[d] & \mathcal{D} \arrow[d, "z_i\cdot"] & \\
& \mathcal{S} & \mathcal{D} \arrow[l, "\psi^{-1}"'] & \\
\big(f(\alpha - e_i)\big)_{\substack{\alpha \in \mathbb{N}^m \\ \alpha_i \neq 0}} & & & \sum_{\substack{\alpha \in \mathbb{N}^m \\ \alpha_i \neq 0}} f(\alpha - e_i) z^\alpha \arrow[lll, mapsto]
\end{tikzcd}$$

Here $e_i \in \mathbb{N}^m$ denotes the ith standard basis vector.

Throughout this paper a code is defined to be an $\mathbb{F}$-linear subspace of some $\mathcal{S}^n$ which is invariant under the forward shifts along all axes. By virtue of the above diagram this can simply be phrased as

Definition 2.2 A linear m-dimensional convolutional code (for short, m-D code) of length n over $\mathbb{F}$ is a $\mathcal{D}$-submodule of $\mathcal{D}^n$. An element of a code is said to be a codeword.

Remark 2.3 In the coding literature (see e.g. [6]) convolutional codes are usually not restricted to sequence spaces whose elements have finite support. There is however no engineering reason behind this. After all every transmitted message created by mankind did have finite length. Convolutional codes with finite support were first studied by Fornasini and Valcher [1, 2, 11]. These authors did define a convolutional code as a submodule of $\tilde{\mathcal{D}}^n$ where $\tilde{\mathcal{D}}$ represents the ring of Laurent polynomials $\mathbb{F}[z_1, \ldots, z_m, z_1^{-1}, \ldots, z_m^{-1}]$. In doing so a convolutional code then corresponds to an $\mathbb{F}$-linear subspace of some $\tilde{\mathcal{S}}^n$, where $\tilde{\mathcal{S}} = \{f : \mathbb{Z}^m \longrightarrow \mathbb{F} \mid f \text{ has finite support}\}$.

Since $\mathcal{D}$ is a Noetherian ring, each code $\mathcal{C} \subseteq \mathcal{D}^n$ is finitely generated. In other words, there exists some $l \in \mathbb{N}$ and a matrix $G \in \mathcal{D}^{n\times l}$ such that $\mathcal{C} = \mathrm{im}_{\mathcal{D}} G$. We call such a matrix G a generator matrix of $\mathcal{C}$. Note that we don't use the row vector notation as common in coding theory. It would force us to use the same notation also for the dual system theoretic version, which is very unusual. The notation $\mathrm{im}_{\mathcal{D}} G$ means of course the set of all Gp with $p \in \mathcal{D}^l$. This notation instead of only $\mathrm{im} G$ will be necessary later when interpreting G as a different type of operator. Analogously, we might also use the notation $\ker_{\mathcal{D}} G = \{p \in \mathcal{D}^l \mid Gp = 0\}$.

As a finitely generated $\mathcal{D}$-module each code has a well-defined rank, say $\mathrm{rank}\, \mathcal{C} = k$. It can simply be calculated as $\mathrm{rank}\, G$, where one may use any generator matrix G of $\mathcal{C}$, considered as a matrix over the quotient field $\mathbb{F}(z_1, \ldots, z_m)$. The rate of $\mathcal{C}$ is defined to be the quotient $\frac{k}{n}$.

The code $\mathcal{C}$ is called free if $\mathcal{C}$ is a free $\mathcal{D}$-module, that is, if $\mathcal{C}$ has a $\mathcal{D}$-basis. This is the case if and only if $\mathcal{C}$ has a generator matrix $G \in \mathcal{D}^{n\times k}$ with $\mathrm{rank}\, G = k = \mathrm{rank}\, \mathcal{C}$. Such a generator matrix is called an encoder. If $\mathcal{C}$ has an encoder, say $G = [G_1, \ldots, G_k] \in \mathcal{D}^{n\times k}$, then each codeword can be written in a unique way as a $\mathcal{D}$-linear combination of $G_1, \ldots, G_k$. This is certainly a very desirable property for a code. It is a well-known fact that each 1-dimensional code, that is each $\mathbb{F}[z_1]$-module is free. However, for higher dimensions, i. e. for $m > 1$, this is not true anymore.

Example 2.4 Let $\mathcal{D} = \mathbb{F}[z_1, z_2]$ and

$$G(z_1, z_2) = \begin{bmatrix} {z_1}^2 & z_1 + z_1 z_2 \\ z_1 & 1 + z_2 \\ z_1 z_2 & z_2 + {z_2}^2 \end{bmatrix}$$

It can easily be shown that $\mathrm{im}_{\mathcal{D}}(G)$ is rate $\frac{1}{3}$ but not free. That is, $\mathrm{im}_{\mathcal{D}}(G)$ has no 3×1 encoder. The code $\mathrm{im}_{\mathcal{D}} \begin{bmatrix} z_1 & 1 & z_2 \end{bmatrix}^{\mathsf{T}}$ is free of rate $\frac{1}{3}$ and properly contains $\mathrm{im}_{\mathcal{D}}(G)$.

It is easy to see that encoder matrices for a given code are unique up to unimodular right multiplication, i. e. for $G_i \in \mathcal{D}^{n\times k}$ with $\operatorname{rank} G_i = k$ it is

$$\mathrm{im}_{\mathcal{D}} G_1 = \mathrm{im}_{\mathcal{D}} G_2 \Longleftrightarrow G_2 = G_1 U \text{ for some } U \in Gl_k(\mathcal{D}). \tag{2.1}$$

An important measure for the 'goodness' of a (convolutional) code is its distance. In the remainder of this section we will introduce this parameter.

Definition 2.5 Let $a \in \mathbb{F}^n$. The *weight* of a is given by the number of nonzero entries of a. It is denoted by $\mathrm{wt}(a)$.

For $w = \sum_{\alpha\in\mathbb{N}^m} b_\alpha z^\alpha \in \mathcal{D}^n$ with $b_\alpha \in \mathbb{F}^n$ the *weight* of w is defined as

$$\mathrm{wt}(w) = \sum_{\alpha\in\mathbb{N}^m} \mathrm{wt}(b_\alpha)\,.$$

Hence the weight of a vector in $\mathcal{D}^n$ measures the distance to the all zero vector by counting all non-zero terms in the vector. The weight has the characteristic of a discrete norm. In particular the weight induces a metric on $\mathcal{D}^n$:

Definition 2.6 Given two elements $w, \tilde{w} \in \mathcal{D}^n$ the *(Hamming) distance* between w and $\tilde{w}$ is given by $\mathrm{dist}(w, \tilde{w}) = \mathrm{wt}(w - \tilde{w})$. Given any m-*D* code $\mathcal{C}$ of length n, the *distance of* $\mathcal{C}$ is defined as

$$\mathrm{dist}(\mathcal{C}) = \min\{\mathrm{dist}(w, \tilde{w}) : w, \tilde{w} \in \mathcal{C}, w \neq \tilde{w}\}.$$

Remark 2.7

i) The Hamming distance defines a metric on $\mathcal{D}^n$ called the Hamming metric.

ii) For a code $\mathcal{C}$ we have that $\mathrm{dist}(\mathcal{C}) = \min\{\mathrm{wt}(w) : w \in \mathcal{C}, w \neq 0\}$. This is because $\mathrm{dist}(w_1, w_2) = \mathrm{wt}(w_1 - w_2)$ and $w_1 - w_2 \in \mathcal{C}$ whenever $w_1, w_2 \in \mathcal{C}$.

The following result is standard in coding theory. It follows immediately from the definition of Hamming distance and from the properties of a metric space.

Lemma 2.8 *Let $\mathcal{C}$ be a convolutional code with $d = \mathrm{dist}(\mathcal{C})$. Let $t = \lfloor \frac{d-1}{2} \rfloor$ where $\lfloor x \rfloor$ denotes the greatest integer that is less than or equal to x. Let $y \in \mathcal{D}^n$. If $w \in \mathcal{C}$ is a codeword such that $\mathrm{dist}(w, y) \leq t$, then w is the unique codeword nearest (with respect to the Hamming metric) to y.*

We say that $\mathcal{C}$ can correct up to t errors. In practice it is not a simple task to compute the transmitted vector $w \in \mathcal{D}^n$ from the received vector $y \in \mathcal{D}^n$. One way to do this is by syndrome decoding. In order to explain this we first state

Proposition 2.9 *Suppose $\mathcal{C}$ is a free convolutional code of rate $\frac{k}{n}$ with encoder $G \in \mathcal{D}^{n\times k}$. Then*

$$0 \longrightarrow \mathcal{D}^k \xrightarrow{G\cdot} \mathcal{D}^n \xrightarrow{\pi} \mathcal{D}^n/\mathcal{C} \longrightarrow 0$$

is a short exact sequence.

The proof is left to the reader.

In the above short exact sequence π is often called a syndrome former. Syndrome decoding works as follows. If $y \in \mathcal{D}^n$ is received, one seeks the vector $e \in \pi^{-1}\big(\pi(y)\big) = y + \mathcal{C}$ of smallest possible weight. The vector y is then decoded as $w := y - e$.

3 Duality between Codes and Behaviors

There have been several instances in the recent literature about coding theory, in which certain types of duality between convolutional codes and behaviors in the system theoretical sense of [13] have been mentioned or used, see, e. g. [3, 8, 11, 12].

In this section we are going to make this duality precise by introducing the appropriate bilinear form. Exploiting the very comprehensive and powerful paper of Oberst [5], quite a lot of results about this duality are available even in the multidimensional case. However, as it seems to us, most interesting is the duality in the 1-dimensional case, where various minimal first-order representations exist and have been studied systematically and exhaustively by Kuijper [4]. They can be translated into corresponding descriptions for codes. This will be studied in Section 4.

We will introduce the notations and the setting along the lines of [5]. Only those results needed for our purposes will be cited afterwards.

Throughout this section the field $\mathbb{F}$ need not be finite; the results hold for any field. First we have to define the underlying setting for the behaviors. Let

$$\mathcal{A} := \mathbb{F}[\![z]\!] = \Big\{ \sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha \;\Big|\; f_\alpha \in \mathbb{F} \Big\}$$

be the set of power series in the m variables $z_1, \ldots, z_m$ over $\mathbb{F}$. On $\mathcal{A}$ we consider the backward shifts along the ith axis followed by truncation; that is, for each $i = 1, \ldots, m$ define

$$\begin{array}{rccc} L_i : & \mathcal{A} & \longrightarrow & \mathcal{A} \\ & \displaystyle\sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha & \longmapsto & \displaystyle\sum_{\alpha\in\mathbb{N}^m} f_{\alpha+e_i} z^\alpha \end{array} \tag{3.1}$$

(see also [5, p. 15]). This action can also be expressed in the following ways

$$L_i\Big(\sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha\Big) = z_i^{-1}\Big(\sum_{\substack{\alpha\in\mathbb{N}^m\\ \alpha_i\neq 0}} f_\alpha z^\alpha\Big) = \Pi_+\Big(z_i^{-1}\sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha\Big) \tag{3.2}$$

where Π_+ denotes the projection which cuts off the terms with negative exponents. Clearly, the operators L_i are $\mathbb{F}$-linear. Moreover, $\mathcal{A}$ gets the structure of a $\mathcal{D}$-module via the scalar multiplication

$$p(z_1, \dots, z_m) \cdot f := p(L_1, \dots, L_m)(f) \in \mathcal{A} \text{ for } p \in \mathcal{D},\ f \in \mathcal{A}.$$

Example 3.1 Let $m = 2$ and $p = 1 + z_1^2 + z_1 z_2 \in \mathcal{D} = \mathbb{F}_5[z_1, z_2]$. Then $p \cdot (1 + 3z_1 z_2^3 + 2z_1^2 + 4z_2) = 3 + 3z_1 z_2^3 + 2z_1^2 + 4z_2 + 3z_2^2$.

The example shows that the notation $p \cdot f$ has to be read with care. It is not the usual convolutional product in $\mathcal{A}$. Instead from (3.2) one can derive the formula

$$\sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha z^\alpha \cdot \sum_{\beta\in\mathbb{N}^m} f_\beta z^\beta = \sum_{\beta\in\mathbb{N}^m} \Big(\sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha f_{\alpha+\beta} \Big) z^\beta. \tag{3.3}$$

Since we never use ordinary convolution in $\mathcal{A}$, this should not cause a confusion.

Remark 3.2 (a) Obviously, $\mathcal{D}$ is a $\mathcal{D}$-submodule of $\mathcal{A}$. However, it is worth mentioning that the canonical injection $\iota : \mathcal{D} \to \mathcal{A},\ p \mapsto p$ is not $\mathcal{D}$-linear. In fact, e. g., $z_1 = \iota(z_1) \neq z_1 \cdot \iota(1) = z_1 \cdot 1 = 0$ in $\mathcal{A}$. This is not really an issue as the inclusion $\mathcal{D} \subset \mathcal{A}$ is never considered in this setting. While $\mathcal{D}$ is the set of operators, either generator matrices for codes or shift operators, $\mathcal{A}$ serves as the space of trajectories for the behaviors.

(b) $\mathcal{A}$ is not finitely generated as $\mathcal{D}$-module, see [5, p. 55].

Each polynomial matrix $G \in \mathcal{D}^{k\times n}$ gives rise to a linear partial difference operator $G(L_1, \dots, L_m)$ which we will denote for short by G, thus

$$\begin{aligned} G :\ \mathcal{A}^n &\longrightarrow \mathcal{A}^k \\ a &\longmapsto G \cdot a := G(L_1, \dots, L_m)(a) \end{aligned}$$

These operators are going to be the objects dual to generator matrices for codes. The following notations will be useful in the sequel. For $G \in \mathcal{D}^{k\times n}$ define

$$\mathrm{im}_{\mathcal{A}} G = \{G \cdot a \mid a \in \mathcal{A}^n\} \subseteq \mathcal{A}^k \quad \text{and} \quad \ker_{\mathcal{A}} G = \{a \in \mathcal{A}^n \mid G \cdot a = 0\} \subseteq \mathcal{A}^n.$$

Definition 3.3 An m-dimensional behavior $\mathcal{B}$ in $\mathcal{A}^n$ is defined to be a $\mathcal{D}$-submodule $\mathcal{B} \subseteq \mathcal{A}^n$ of the form $\mathcal{B} = \ker_{\mathcal{A}} G$ for some $G \in \mathcal{D}^{k\times n}$ (not necessarily of full row rank).

This setting is identical to the study of m-D-discrete-time systems in the behavioral context, see e. g. [7].

We observe that, while each $\mathcal{D}$-submodule of $\mathcal{D}^n$ is a code, not every $\mathcal{D}$-submodule of $\mathcal{A}^n$ is a behavior. Characterizations for an $\mathbb{F}$-subspace of $\mathcal{A}^n$ being a behavior are given in the 1-dimensional case in [13, III.1] and for the general case in [5, p. 61/62].

Now the bilinear form to be used for the duality is obvious. For each $n \geq 1$ a $\mathcal{D}$-bilinear non-degenerate form is given by (cf. [5, p. 22])

$$\begin{aligned} \mathcal{D}^n \times \mathcal{A}^n &\longrightarrow \quad \mathcal{A} \\ (p, a) \quad &\longmapsto \langle p, a\rangle := p^{\mathsf{T}} \cdot a = \sum_{i=1}^{n} p_i \cdot a_i = \sum_{i=1}^{n} p_i(L_1, \ldots, L_m)(a_i) \end{aligned} \tag{3.4}$$

where $p = (p_1, \ldots, p_n)^{\mathsf{T}}$, $a = (a_1, \ldots, a_n)^{\mathsf{T}}$. In the literature related to codes and behaviors also a certain $\mathbb{F}$-bilinear form has been used, see [8] and [12, p. 20]. We will clarify the relationship between this one and (3.4) at the end of this section.

Using the above bilinear form we define the duals in the obvious way.

Definition 3.4 (a) The dual of a subset $\mathcal{B} \subseteq \mathcal{A}^n$ is defined to be $\mathcal{B}^{\perp} := \{p \in \mathcal{D}^n \mid \langle p, a\rangle = 0 \text{ for all } a \in \mathcal{B}\}$.

(b) The dual of a subset $\mathcal{C} \subseteq \mathcal{D}^n$ is given by $\mathcal{C}^{\perp} := \{a \in \mathcal{A}^n \mid \langle p, a\rangle = 0 \text{ for all } p \in \mathcal{C}\}$.

Obviously, duals are $\mathcal{D}$-modules and one has $\mathcal{B} \subseteq \mathcal{B}^{\perp\perp}$ as well as $\mathcal{C} \subseteq \mathcal{C}^{\perp\perp}$.

Now we are in the position to state the results given in [5]. Essentially, they amount to the fact that $\mathcal{A}$ is a large injective cogenerator in the category of $\mathcal{D}$-modules. Instead of going into an explanation of this statement, we will simply extract from [5] the following consequences of this very strong result. Statements (4), (5), and (7) of the next theorem are exactly the duality between codes and behaviors we were looking for.

Theorem 3.5 *Let* $P \in \mathcal{D}^{l\times n}$, $Q \in \mathcal{D}^{k\times l}$, $R \in \mathcal{D}^{r\times n}$. *Then*

(1) *If the sequence* $\mathcal{D}^k \xrightarrow{Q^{\mathsf{T}}} \mathcal{D}^l \xrightarrow{P^{\mathsf{T}}} \mathcal{D}^n$ *is exact, then so is the sequence* $\mathcal{A}^n \xrightarrow{P} \mathcal{A}^l \xrightarrow{Q} \mathcal{A}^k$.

(2) $\ker_{\mathcal{A}} P \subseteq \ker_{\mathcal{A}} R$ *if and only if* $R = XP$ *for some* $X \in \mathcal{D}^{r\times l}$.

(3) *If* $\operatorname{rank} P = l$, *then the operator* $P : \mathcal{A}^n \to \mathcal{A}^l$ *is surjective.*

(4) $\big(\mathrm{im}_{\mathcal{D}} Q^{\mathsf{T}}\big)^{\perp} = \ker_{\mathcal{A}} Q$.

(5) $\big(\ker_{\mathcal{A}} Q\big)^{\perp} = \mathrm{im}_{\mathcal{D}} Q^{\mathsf{T}}$.

(6) $(\mathrm{im}_{\mathcal{A}} Q)^{\perp} = \ker_{\mathcal{D}} Q^{\mathsf{T}}$.

(7) $\mathcal{C} = \mathcal{C}^{\perp\perp}$ *and* $\mathcal{B} = \mathcal{B}^{\perp\perp}$ *for each code* $\mathcal{C} \in \mathcal{D}^n$ *and each behavior* $\mathcal{B} \in \mathcal{A}^n$.

(3) means in other words, for each P with full row rank and for each $g \in \mathcal{A}^l$ the associated linear partial difference equation $P \cdot f = g$ has a solution in $\mathcal{A}^n$. This is a well-known fact in the 1-dimensional case, that is, $\mathcal{D} = \mathbb{F}[z_1]$. Even more, one can also prescribe initial conditions up to a certain order. In the m-dimensional case this is more involved. Statement (4) shows especially that the dual of a code is not only a $\mathcal{D}$-module but even a behavior.

As for the proof, all the above results go back to [5, p. 33], which is just the large injective cogenerator property. However, we will give some more detailed references and arguments from the paper to show how things are related with each other, although this might be a bit different from the order they have been proven.

(1) is exactly the injectivity of the module $\mathcal{A}$ which is defined at [5, p. 24]. (2) is at [5, p. 36]. (3) is a consequence of (1). (4) and (5) are at [5, p. 30/31], but they can also be derived directly from the above as follows. (4) and also (6) follow immediately from

$$\langle Q^{\mathsf{T}} p, a\rangle = \langle p, Q \cdot a\rangle \text{ for each } p \in \mathcal{D}^k \text{ and } a \in \mathcal{A}^l \tag{3.5}$$

together with the non-degeneracy of the bilinear form (3.4). (5) can be shown with the help of (2) via

$$\begin{aligned} p \in (\ker_{\mathcal{A}} Q)^{\perp} \Longleftrightarrow \ker_{\mathcal{A}} Q \subseteq \ker_{\mathcal{A}} p^{\mathsf{T}} &\Longleftrightarrow p^{\mathsf{T}} = v^{\mathsf{T}} Q \text{ for some } v \in \mathcal{D}^k \\ &\Longleftrightarrow p \in \mathrm{im}_{\mathcal{D}} Q^{\mathsf{T}}. \end{aligned}$$

(7) is a consequence from (4) and (5).

Remark 3.6 (Compare with Remark 2.3). If convolutional codes are defined as submodules of $\tilde{\mathcal{D}}^n$, where $\tilde{\mathcal{D}}$ represents the ring of Laurent polynomials $\mathbb{F}[z, z^{-1}]$ then this results in a duality between codes and linear behaviors defined on $\tilde{\mathcal{A}}^n$, where $\tilde{\mathcal{A}} := \mathbb{F}[\![z, z^{-1}]\!]$ is the ring of formal power series in the variables $z_1, \ldots, z_m, z_1^{-1}, \ldots, z_m^{-1}$.

Next we want to concentrate on two specific descriptions of behaviors. They will be of significance for 1-dimensional first-order-representations in the next section. In fact, the following two types of representations, applicable to both, codes and behaviors, are dual to each other as will be proven next. They specialize to the so-called (P, Q, R)- and (K, L, M)-representations in the 1-dimensional case.

Theorem 3.7 *Let $R \in \mathcal{D}^{n\times l}$, $N \in \mathcal{D}^{k\times l}$, and $M \in \mathcal{D}^{k\times n}$. Then the following are true.*

(a) *The module $R \cdot (\ker_{\mathcal{A}} N) := \{R \cdot \zeta \mid \zeta \in \mathcal{A}^l,\ N \cdot \zeta = 0\} \subseteq \mathcal{A}^n$ is a behavior and its dual is given by $\big(R \cdot (\ker_{\mathcal{A}} N)\big)^{\perp} = \{p \in \mathcal{D}^n \mid R^{\mathsf{T}} p \in \mathrm{im}_{\mathcal{D}} N^{\mathsf{T}}\}$.*

(b) *The module $\{a \in \mathcal{A}^n \mid M \cdot a \in \mathrm{im}_{\mathcal{A}} N\} \subseteq \mathcal{A}^n$ is a behavior. Its dual is $\{a \in \mathcal{A}^n \mid M \cdot a \in \mathrm{im}_{\mathcal{A}} N\}^{\perp} = M^{\mathsf{T}}(\ker_{\mathcal{D}} N^{\mathsf{T}})$.*

(c) $\big(R(\ker_{\mathcal{D}} N)\big)^{\perp} = \{a \in \mathcal{A}^n \mid R^{\mathsf{T}} \cdot a \in \mathrm{im}_{\mathcal{A}} N^{\mathsf{T}}\}$.

(d) $\{p \in \mathcal{D}^n \mid Mp \in \mathrm{im}_{\mathcal{D}} N\}^{\perp} = M^{\mathsf{T}} \cdot (\ker_{\mathcal{A}} N^{\mathsf{T}})$.

Proof: (a) The first part is proven in [5, p. 26]. As for the second part, note the following equivalences, which hold for each $p \in \mathcal{D}^n$ using equation (3.5)

$$\begin{aligned}\langle p, R \cdot a\rangle = 0 \,\forall\, a \in \ker_{\mathcal{A}} N &\iff \langle R^{\mathsf{T}} p, a\rangle = 0 \,\forall\, a \in \ker_{\mathcal{A}} N \\ &\iff R^{\mathsf{T}} p \in (\ker_{\mathcal{A}} N)^{\perp} = \mathrm{im}_{\mathcal{D}} N^{\mathsf{T}}.\end{aligned}$$

(b) Using $N = 0$ in (a) we obtain especially that a $\mathcal{D}$-submodule of the form $\mathrm{im}_{\mathcal{A}} R$ is a behavior. Thus, write $\mathrm{im}_{\mathcal{A}} N = \ker_{\mathcal{A}} Q$ with some appropriate $Q \in \mathcal{D}^{q \times k}$. Then $\{a \in \mathcal{A}^n \mid M \cdot a \in \mathrm{im}_{\mathcal{A}} N\} = \ker_{\mathcal{A}} QM$ is a behavior (see also [5, p. 27]) and moreover

$$\begin{aligned}\{a \in \mathcal{A}^n \mid M \cdot a \in \mathrm{im}_{\mathcal{A}} N\}^{\perp} &= (\ker_{\mathcal{A}} QM)^{\perp} = \mathrm{im}_{\mathcal{D}} (QM)^{\mathsf{T}} = M^{\mathsf{T}}(\mathrm{im}_{\mathcal{D}} Q^{\mathsf{T}}) \\ &= M^{\mathsf{T}} \cdot \big((\ker_{\mathcal{A}} Q)^{\perp}\big) = M^{\mathsf{T}}\big((\mathrm{im}_{\mathcal{A}} N)^{\perp}\big) = M^{\mathsf{T}}(\ker_{\mathcal{D}} N^{\mathsf{T}}).\end{aligned}$$

(c) and (d) follow now from (a) and (b) with Thm. 3.5 (7). □

In the following we want to briefly discuss parity check matrices for multidimensional codes.

Definition 3.8 Let $\mathcal{C} \subseteq \mathcal{D}^n$ be a code. A matrix $H \in \mathcal{D}^{l \times n}$ is called a parity check matrix of $\mathcal{C}$ if $\mathcal{C} = \ker_{\mathcal{D}} H$.

Not each code has a parity check matrix; e. g. for $\mathcal{D} = \mathbb{F}[z_1]$ the code $\mathrm{im}_{\mathcal{D}} \begin{bmatrix} z_1 \\ z_1 \end{bmatrix}$ has no parity check matrix, since each matrix $H \in \mathbb{F}[z_1]^{l \times 2}$ having $(z_1, z_1)^{\mathsf{T}}$ in its kernel, would also have $(1,1)^{\mathsf{T}} \in \ker_{\mathcal{D}} H$.

The following result about the existence of parity check matrices can be found in [12, 3.3.8].

Theorem 3.9 *Let $\mathcal{C} = \mathrm{im}_{\mathcal{D}} G$ with $G \in \mathcal{D}^{n \times k}$ be a free code, thus $\mathrm{rank}\, G = k$. Then $\mathcal{C}$ has a parity check matrix if and only if G is minor-prime, that is, if the greatest common divisor of all full-size minors of G is a unit in $\mathcal{D}$. If a parity check matrix exists, then one also has a parity check matrix $H \in \mathcal{D}^{(n-k) \times n}$ with $\mathrm{rank}\, H = n - k$.*

This result can be dualized by use of Thm. 3.5.

Theorem 3.10 *Let $\mathcal{C} \subseteq \mathcal{D}^n$ be a free code. Then $\mathcal{C}$ has a parity check matrix if and only if the behavior $\mathcal{C}^{\perp} \subseteq \mathcal{A}^n$ has an image-representation, i. e.*

$$\mathcal{C} = \ker_{\mathcal{D}} H \text{ for some } H \in \mathcal{D}^{l \times n} \iff \mathcal{C}^{\perp} = \mathrm{im}_{\mathcal{A}} H^{\mathsf{T}} \text{ for some } H \in \mathcal{D}^{l \times n}.$$

Hence a behavior $\ker_{\mathcal{A}} G \subseteq \mathcal{A}^n$ has an image-representation if and only if G is minor-prime.

Proof: follows from Thm. 3.5 (1), (6), and (7). □

Recall that for 1-dimensional behaviors the existence of image-representations is equivalent to controllability, see [13]. For $m > 1$, at least one direction is true, namely, behaviors with image-representations are always controllable, see [14, Thm. 4.2]. Equivalence can be established for $m = 2$ or for $m \geq 2$ if certain directions of the time-space axes are two-sided, see [7] and [15, Thm. 6].

At the end of this section we want to discuss the relationship of the above bilinear form with an $\mathbb{F}$-bilinear form which has been used as well in the literature within this context. Let

$$\begin{array}{ccc} \mathcal{D}^n \times \mathcal{A}^n & \longrightarrow & \mathbb{F} \\ \Big(\sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha z^\alpha, \sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha\Big) & \longmapsto & \langle\!\langle p, f\rangle\!\rangle := \sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha^\mathsf{T} f_\alpha \end{array}$$

where $p_\alpha^\mathsf{T} f_\alpha \in \mathbb{F}$ denotes the usual scalar product in $\mathbb{F}^n$. Observe that the sum on the right hand side is indeed finite.

Example 3.11 Let $\mathcal{D} = \mathbb{F}_2[z_1]$ and $n = 2$. For $p = (1,0)^\mathsf{T} \in \mathcal{D}^2$ and $f = (z_1, 1)^\mathsf{T} \in \mathcal{A}^2$ we obtain $\langle\!\langle p, f\rangle\!\rangle = (1,0)\begin{pmatrix}0\\1\end{pmatrix} = 0$, whereas the previously used $\mathcal{D}$-bilinear form yields $\langle p, f\rangle = (1,0)\begin{pmatrix}z_1\\1\end{pmatrix} = z_1$. Hence p and f are orthogonal with respect to $\langle\!\langle\ ,\ \rangle\!\rangle$ but not with respect to $\langle\ ,\ \rangle$.

However, there is a close relationship between these two forms as we will derive next. In order to do so, we use the notation $L^\alpha := L_1^{\alpha_1} \circ \ldots \circ L_m^{\alpha_m}$ for $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$ and the shifts L_i defined in (3.1). Let $p = \sum'_{\alpha\in\mathbb{N}^m} p_\alpha z^\alpha \in \mathcal{D}^n$ and $f = \sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha \in \mathcal{A}^n$. Firstly, using the very definition (3.4) and equation (3.3) one obtains

$$\begin{aligned} \langle p, f\rangle = 0 &\iff \sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha^\mathsf{T} L^\alpha(f) = 0 \iff \sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha^\mathsf{T} \sum_{\beta\in\mathbb{N}^m} f_{\beta+\alpha} z^\beta = 0 \\ &\iff \sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha^\mathsf{T} f_{\beta+\alpha} = 0\ \forall\, \beta \in \mathbb{N}^m \iff \langle\!\langle p, z^\beta \cdot f\rangle\!\rangle = 0\ \forall\, \beta \in \mathbb{N}^m. \end{aligned}$$

Secondly, it is

$$\langle\!\langle z^\beta p, f\rangle\!\rangle = \langle\!\langle \sum{}'_{\alpha\in\mathbb{N}^m} p_\alpha z^{\alpha+\beta}, \textstyle\sum_{\alpha\in\mathbb{N}^m} f_\alpha z^\alpha\rangle\!\rangle = \displaystyle\sum_{\alpha\in\mathbb{N}^m}{}' p_\alpha^\mathsf{T} f_{\alpha+\beta} = \langle\!\langle p, z^\beta \cdot f\rangle\!\rangle$$

by virtue of (3.3).

These two observations lead to the fact that both bilinear forms yield the same duals for $\mathcal{D}$-submodules of $\mathcal{A}^n$ or $\mathcal{D}^n$. Indeed, if $\mathcal{B} \subseteq \mathcal{A}^n$ is a $\mathcal{D}$-submodule, then

$$\{p \in \mathcal{D}^n \mid \langle\langle p, f\rangle\rangle = 0 \;\forall\, f \in \mathcal{B}\} = \{p \in \mathcal{D}^n \mid \langle\langle p, z^\beta \cdot f\rangle\rangle = 0 \;\forall\, f \in \mathcal{B} \;\forall\, \beta \in \mathbb{N}^m\} = \mathcal{B}^\perp.$$

Similarly, for a $\mathcal{D}$-submodule $\mathcal{C} \subseteq \mathcal{D}^n$ one obtains

$$\begin{aligned}\{f \in \mathcal{A}^n \mid \langle\langle p, f\rangle\rangle = 0 \;\forall\, p \in \mathcal{C}\} &= \{f \in \mathcal{A}^n \mid \langle\langle z^\beta p, f\rangle\rangle = 0 \;\forall\, p \in \mathcal{C} \;\forall\, \beta \in \mathbb{N}^m\}\\ &= \{f \in \mathcal{A}^n \mid \langle\langle p, z^\beta \cdot f\rangle\rangle = 0 \;\forall\, p \in \mathcal{C} \;\forall\, \beta \in \mathbb{N}^m\} = \mathcal{C}^\perp.\end{aligned}$$

4 First-Order Representations for 1-Dimensional Codes

In this last section we restrict to the 1-dimensional case, thus $\mathcal{D} = \mathbb{F}[z]$ denotes the polynomial ring in one variable over $\mathbb{F}$ and each submodule $\mathcal{C} \in \mathcal{D}^n$ is a convolutional code in the sense of, e. g., [6]. Using the duality results from the last section and certain well-studied first-order representations for behaviors, we can derive analogous descriptions for codes along with minimality and uniqueness results.

The main source for this section is the book [4] about behaviors. Although [4] deals with the field $\mathbb{R}$, it can be checked that the results hold true for any field.

We need to introduce the following parameter, called degree, for 1-dimensional codes. It is the analogue to the McMillan degree or order of a system. Let $\mathcal{C} = \mathrm{im}_{\mathcal{D}} G$ with $G \in \mathcal{D}^{n\times k}$ and $\operatorname{rank} G = k$, a non-restrictive assumption. The degree $\delta(\mathcal{C})$ is defined to be the maximum degree of all $k \times k$-minors of G. The degree is sometimes also called the complexity of the code $\mathcal{C}$ (see [6, 2.7]) and it corresponds to the McMillan degree of the associated behavior under the duality studied in the last section, see Thm 3.5 (4) and [13, p. 276]. Equation (2.1) shows that the degree does not depend on the choice of the encoder G. A code of degree $\delta(\mathcal{C}) = 0$ is in essence a block code.

Theorem 4.1 *Let $\mathcal{C} = \mathrm{im}_{\mathcal{D}} G$ with $G \in \mathcal{D}^{n\times k}$ be a rate $\frac{k}{n}$ code of degree $\delta(\mathcal{C}) = \delta > 0$.*

(a) *There exist matrices $(P, Q, R) \in \mathbb{F}^{\delta\times(\delta+k)} \times \mathbb{F}^{\delta\times(\delta+k)} \times \mathbb{F}^{n\times(\delta+k)}$ such that*

$$\mathcal{C} = R\big(\ker_{\mathcal{D}}(zP + Q)\big).$$

Moreover,
(i) $\operatorname{rank} P = \delta$,
(ii) $\operatorname{rank} \begin{bmatrix} P \\ R \end{bmatrix} = \delta + k$,
(iii) $zP + Q \in \mathcal{D}^{\delta\times(\delta+k)}$ *is left-prime.*

(b) *If* $\mathcal{C} = R\big(\ker_{\mathcal{D}}(zP+Q)\big) = \tilde{R}\big(\ker_{\mathcal{D}}(z\tilde{P}+\tilde{Q})\big)$ *with matrix triples* (P,Q,R) *and* $(\tilde{P},\tilde{Q},\tilde{R})$ *being of the sizes as in (a), then*

$$(\tilde{P},\tilde{Q},\tilde{R}) = (T^{-1}PS, T^{-1}QS, RS) \text{ for some } T \in Gl_{\delta}(\mathbb{F}) \text{ and } S \in Gl_{\delta+k}(\mathbb{F}).$$

Proof: (a) By Thm. 3.5 (4) we have $\mathcal{C}^{\perp} = \ker_{\mathcal{A}} G^{\mathsf{T}}$. Without loss of generality we may assume that G is column-reduced, that is, δ is the sum of the column degrees of G. From [4, 5.17] we obtain matrices $(K,L,M) \in \mathbb{F}^{(\delta+k)\times\delta} \times \mathbb{F}^{(\delta+k)\times\delta} \times \mathbb{F}^{(\delta+k)\times n}$ such that $\mathcal{C}^{\perp} = \{a \in \mathcal{A}^n \mid M \cdot a \in \mathrm{im}_{\mathcal{A}}(zK+L)\}$. Indeed, the parameter ord (Σ) in [4, p. 128] is equal to the degree, cf. [4, 3.11 and 2.22]. Setting $(P,Q,R) = (K^{\mathsf{T}}, L^{\mathsf{T}}, M^{\mathsf{T}})$ and using Thm. 3.7 (b) and Thm. 3.5 (7) we obtain the desired representation. Furthermore, [4, 5.17] shows that the triple (K,L,M) is minimal with respect to row and column size of the matrix K (or L). Hence, use of [4, 4.32] leads to (i) – (iii).

(b) follows from [4, 4.40] and Thm. 3.7 (c). $\square$

In fact, the proof shows more. The above given sizes of the matrices (P,Q,R) are minimal among all representations of this type. The minimality is equivalent to the properties (i) – (iii). An alternative direct proof, without using duality, is given in the paper [10].

In exactly the same way we can derive so-called (K,L,M)-representations for codes. For this we use [4, 5.10 and 4.3] and dualize these representations using Thm. 3.7 (a) and (d). This results in [8, Theorem 3.1 and Theorem 3.4]:

Theorem 4.2 *Let* $\mathcal{C} = \mathrm{im}_{\mathcal{D}} G$ *with* $G \in \mathcal{D}^{n\times k}$ *be a rate* $\frac{k}{n}$ *code of degree* $\delta(\mathcal{C}) = \delta > 0$.

(a) *There exist matrices* $(K,L,M) \in \mathbb{F}^{(\delta+n-k)\times\delta} \times \mathbb{F}^{(\delta+n-k)\times\delta} \times \mathbb{F}^{(\delta+n-k)\times n}$ *so that*

$$\mathcal{C} = \{p \in \mathcal{D}^n \mid Mp \in \mathrm{im}_{\mathcal{D}}(zK+L)\}.$$

Moreover,
(i) $\mathrm{rank}\, K = \delta$,
(ii) $\mathrm{rank}\, [K, M] = \delta + n - k$,
(iii) $[zK + L \mid M]$ *is left-prime over the polynomial ring* $\mathcal{D}$.

(b) *If* $\mathcal{C} = \{p \in \mathcal{D}^n \mid Mp \in \mathrm{im}_{\mathcal{D}}(zK+L)\} = \{p \in \mathcal{D}^n \mid \tilde{M}p \in \mathrm{im}_{\mathcal{D}}(z\tilde{K}+\tilde{L})\}$ *with matrix triples* (K,L,M) *and* $(\tilde{K},\tilde{L},\tilde{M})$ *being of the sizes as in (a), then*

$$(\tilde{K},\tilde{L},\tilde{M}) = (T^{-1}KS, T^{-1}LS, T^{-1}M) \text{ for some } T \in Gl_{\delta+n-k}(\mathbb{F}),\ S \in Gl_{\delta}(\mathbb{F}).$$

Generalized first order representations as described in the above two theorems are very useful in the design of convolutional codes with large distance and which can be encoded in an efficient manner. We refer the interested reader to [8, 9].

Conclusion

The paper did show that multidimensional convolutional codes are powerful encoding devices for the transmission of data over a noisy channel. Since these codes are dual objects to multidimensional systems the algebraic theory of linear systems can be fruitfully applied.

Diederich Hinrichsen, to whom this paper is dedicated, contributed over the years significantly to algebraic systems theory. As it happens often in research a contribution in one area bears unexpected fruits in another research field. We believe that the recent cross fertilization between coding theory and systems theory is such an instance.

References

[1] E. Fornasini and M.E. Valcher. Algebraic aspects of 2D convolutional codes. *IEEE Trans. Inform. Theory*, IT-40(4):1068–1082, 1994.

[2] E. Fornasini and M.E. Valcher. Multidimensional systems with finite support behaviors: Signal structure, generation, and detection. *SIAM J. Control Optim.*, 36(2):760–779, 1998.

[3] G. D. Forney and M. D. Trott. Controllability, observability, and duality in behavioral group systems. In *Proc. of the 34th IEEE Conference on Decision and Control*, pages 3259–3264, New Orleans, Louisiana, 1995.

[4] M. Kuijper. *First-Order Representations of Linear Systems.* Birkhäuser, Boston, 1994.

[5] U. Oberst. Multidimensional constant linear systems. *Acta Appl. Math*, 20:1–175, 1990.

[6] Ph. Piret. *Convolutional Codes, an Algebraic Approach.* MIT Press, Cambridge, MA, 1988.

[7] M.P.M Rocha. *Structure and Representation of 2D Systems.* PhD thesis, University of Groningen, 1990.

[8] J. Rosenthal, J. M. Schumacher, and E.V. York. On behaviors and convolutional codes. *IEEE Trans. Inform. Theory*, 42(6):1881–1891, 1996.

[9] J. Rosenthal and E.V. York. BCH convolutional codes. Technical report, University of Notre Dame, Dept. of Mathematics, October 1997. Preprint # 271. Available at http://www.nd.edu/˜rosen/preprints.html.

[10] R. Smarandache, H. Gluesing-Luerssen, and J. Rosenthal. Generalized first order descriptions and canonical forms for convolutional codes. In *Proceedings of the MTNS*, Padova, Italy, 1998. To appear.

[11] M.E. Valcher and E. Fornasini. On 2D finite support convolutional codes: an algebraic approach. *Multidim. Sys. and Sign. Proc.*, 5:231–243, 1994.

[12] P. Weiner. *Multidimensional Convolutional Codes*. PhD thesis, University of Notre Dame, 1998. Available at http://www.nd.edu/˜rosen/preprints.html.

[13] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Automat. Control*, AC-36(3):259–294, 1991.

[14] J. Wood, E. Rogers, and D. H. Owens. Controllable and autonomous nD systems. To appear in *Multidimensional Syst. and Sign. Processing*, 1997.

[15] J. Wood and E. Zerz. Notes on the definition of behavioural controllability. Submitted to *Systems and Control Letters*, 1998.