On 2D finite support convolutional codes: an algebraic approach

Maria Elena Valcher and Ettore Fornasini
Dept. of Electronics and Comp. Sci., University of Padova
via Gradenigo 6/a, 35131 Padova, Italy

Abstract

2D finite codes are defined as families of compact support sequences indexed in $\mathbf{Z} \times \mathbf{Z}$ and taking values in $\mathbf{F}^n$, $\mathbf{F}$ a Galois field. Several properties of encoders, decoders and syndrome decoders are discussed under different hypotheses on the code structure, and related to the injectivity and primeness of the corresponding polynomial matrices in two variables.
Dual codes are finally introduced as families of parity checks on a given modular code, and related to the standard theory of 2D behaviours.

1 Introduction

Since the early seventies, the pioneering work of Forney [1, 2] made it quite clear that the theory of discrete-time multidimensional linear systems over a finite field provides a very convenient setting for the analysis of convolutional codes. On the other hand, in the algebraic contex many questions concerning convolutional codes proved to have answers that seem quite illuminating and useful for systems and control applications.
However, even if both fields exhibit some common research directions and resort to similar mathematical tools, the coding point of view is somewhat different from that of linear systems. Actually, in system theory the interest centers around input-output relations, while in coding theory what is most important is the set of output sequences of the encoder, i.e. the internal structure of the code.
Quite recently, the behavioural approach, developed by J.C.Willems [3] for the analysis of dynamical systems, has been applied to the investigation of 1D and 2D convolutional codes [$4 \div 6$]. This new framework seems to be quite effective in the 2D case, since it allows to investigate the internal properties of the code without explicitly referring to the machinery which underlies the codewords generation and, in particular, without making any assumption on the ordering of two dimensional data. So, in principle, no artificial notion of causality in $\mathbf{Z} \times \mathbf{Z}$, and, consequently, no a priori restriction on the supports of the signals are needed. Indeed, the finite-support constraint we shall introduce in a while on two-dimensional codewords does not follow from causality considerations, but corresponds to the fact that most of 2D information sequences encountered in the applications do not infinitely extend in $\mathbf{Z} \times \mathbf{Z}$.

In this comunication we aim to analyse the algebraic properties of two-dimensional convolutional codes whose codewords have finite support, and discuss how they are related with more general classes of 2D codes, that have been modelled in [6] as 2D complete behaviours.

The paper is organized as follows: in the next section, 2D modular codes are defined and some fundamental requirements on the encoding and the decoding maps, which translate into specific constraints on the algebraic structure of the code, are introduced. As any code can be generated by different encoders, in section 3 we discuss different sets of necessary and sufficient conditions, which guarantee the equivalence of two encoders. The analysis is carried out both in the general case and for specific classes of 2D codes, such as free modular, finite convolutional and finite basic codes.
In the last section, we introduce 2D codes with infinite support (unrestricted 2D behaviours) as suitable algebraic duals of modular codes. In this context, a dual code can be viewed as the space of all parity checks that can be applied to a received sequence to decide whether it belongs to the code.
The existence of a finite set of finite support parity checks for a code $\mathcal{C}$, which allow for an unambiguous identification of its codewords, a (*syndrome decoder*) is shown to depend on both the structure of the dual code and the algebraic properties of the encoders of $\mathcal{C}$.

2 Finite convolutional codes

Let $\mathbf{F}$ be a finite field and denote by $\mathcal{F}_\infty^n$ the set of the sequences indexed on the discrete plane $\mathbf{Z} \times \mathbf{Z}$ and taking values in $\mathbf{F}^n$. In the sequel, it will be convenient to represent the elements of $\mathcal{F}_\infty^n$ via formal power series, by associating any sequence $\mathbf{w} := \{w(h,k)\}$ with the series

$$\sum_{h,k \,\in\, \mathbf{Z}} w(h,k) z_1^h z_2^k. \tag{2.1}$$

To avoid cumbersome notations, we will adopt the same symbol both for a sequence and for the associated power series, and denote the coefficient of $z_1^h z_2^k$ in any series $\mathbf{w}$ as $(\mathbf{w}, z_1^h z_2^k)$. The main advantage in using formal power series is that many linear operators can be represented by appropriate matrices with elements in $\mathcal{F}_\pm := \mathbf{F}[z_1, z_2, z_1^{-1}, z_2^{-1}]$, the ring of 2D Laurent polynomials (L-polynomials). This way the fundamental operator properties find an immediate counterpart in terms of the structure of the corresponding matrices and, in particular, of their factors.

Definition A matrix $G(z_1, z_2) \in \mathcal{F}_\pm^{k \times n}$ is

- $\mathcal{F}_\pm$-*unimodular*, if $k = n$ and $\det G$ is a unit in $\mathcal{F}_\pm$;
- *left factor prime* (ℓFP), if for every factorization $G = T\bar{G}$, with $T(z_1, z_2) \in \mathcal{F}_\pm^{k \times k}$, T is $\mathcal{F}_\pm$-unimodular;
- *left zero prime* (ℓZP), if the ideal generated by the maximal order minors of G is the ring $\mathcal{F}_\pm$ itself.

A 2D *code of length* n over $\mathbf{F}$ is any subset of $\mathcal{F}_\infty^n$. In this paper we will mostly deal with *finite codes*, i.e. subsets of $\mathcal{F}_\infty^n$ whose elements have finite support. By the bijective correspondence between sequences indexed in $\mathbf{Z}\times\mathbf{Z}$ and formal power series, we identify each compact support sequence with an element of $\mathcal{F}_\pm^n$,the $\mathcal{F}_\pm$-module of n-dimensional row vectors with entries in $\mathcal{F}_\pm$. Accordingly, a 2D finite code $\mathcal{C}$ of length n is defined as a subset of $\mathcal{F}_\pm^n$.

In order to introduce a convolutional structure on $\mathcal{C}$, the set of its sequences has to be endowed with some properties, which constitute the mathematical formalization of very natural requirements. The most common ones are linearity and shift-invariance, i.e. the closure of $\mathcal{C}$ under shift and superposition.

(a) [Linearity] If $\mathbf{w}_1$ and $\mathbf{w}_2$ belong to $\mathcal{C}$, then $\alpha\mathbf{w}_1+\beta\mathbf{w}_2$ belongs to $\mathcal{C}$ for every α and β in $\mathbf{F}$.

(b) [Shift-Invariance] $\mathbf{w}\in\mathcal{C}$ implies that $\mathbf{v}=z_1^h z_2^k\mathbf{w}\in\mathcal{C}$ for every $h,k\in\mathbf{Z}$, i.e. $\mathcal{C}$ is invariant w.r.t. the shifts in $\mathbf{Z}\times\mathbf{Z}$ along the coordinate axes.

Codes with properties (a) and (b) can be characterized as $\mathcal{F}_\pm$-submodules of $\mathcal{F}_\pm^n$, and will be called *modular codes*. Moreover, as $\mathcal{F}_\pm^n$ is an $\mathcal{F}_\pm$-Noetherian module [7], $\mathcal{C}$ is finitely generated, i.e. there exists a finite set of row vectors $\mathbf{g}_1,\mathbf{g}_2,...,\mathbf{g}_k$ in $\mathcal{F}_\pm^n$ such that

$$\mathcal{C}=\{\sum_{i=1}^{k}u_i\mathbf{g}_i : u_i\in\mathcal{F}_\pm\}=\{\mathbf{u}G : \mathbf{u}\in\mathcal{F}_\pm^k\}=:\mathrm{Im}_\pm G, \tag{2.2}$$

where $G(z_1,z_2)$ denotes the L-polynomial matrix $G=\mathrm{col}\{\mathbf{g}_1,\mathbf{g}_2,...,\mathbf{g}_k\}$.

Once a family of generators has been chosen, the matrix G constitutes an encoder, which generates all the codewords of $\mathcal{C}$ as the information sequence $\mathbf{u}$ varies over $\mathcal{F}_\pm^k$. It may happen, however, that different information sequences in $\mathcal{F}_\pm^k$ produce the same codeword, thus resulting indistinguishable at the decoding stage. Such a drawback can be avoided if and only if G induces an injective map or, equivalently, has full row rank over the field of rational functions $\mathbf{F}(z_1,z_2)$. Since there exist submodules of $\mathcal{F}_\pm^n$ which are not free, not every modular code admits an injective encoder. Finite codes which are free $\mathcal{F}_\pm$-modules are called *free modular codes*.

Example Let $\mathbf{F}=GF(2)$ and consider the modular codes generated by the following encoders

$$\begin{aligned}
G_1(z_1,z_2) &:= \begin{bmatrix} 1 & z_1 & z_2 \\ z_1^{-1} & 1 & z_2 \end{bmatrix} \\
G_2(z_1,z_2) &:= \begin{bmatrix} 1 & z_1 & z_2 \\ z_1 & z_2+z_1 & 1 \\ z_1+1 & z_2 & z_2+1 \end{bmatrix} \\
G_3(z_1,z_2) &:= \begin{bmatrix} (z_1+1)(z_2+1) & (z_1+1)\,z_1z_2^{-1} \\ (z_1+z_2^{-1})(z_2+1) & (z_1+z_2^{-1})\,z_1z_2^{-1} \end{bmatrix}.
\end{aligned}$$

As G_1 is full row rank, the code generated by G_1 is a free module.

Even though G_2 is not full rank, the code it generates is free, because the sum of the first two rows of G_2 gives the third one.
Finally, $\mathrm{Im}_{\pm}G_3$ is not free. Actually, G_3 is not full rank, so if there were a basis, it would consist of a single row vector $\mathbf{p} \in \mathcal{F}_{\pm}^2$. The rows of G_3, being elements of the module generated by $\mathbf{p}$, should be L-polynomial multiples of $\mathbf{p}$, and this requirement determines $\mathbf{p}$ as $\mathbf{p} = [z_2 + 1 \quad z_1 z_2^{-1}]$, modulo a unit factor $z_1^n z_2^m$, $n, m \in \mathbf{Z}$. Such a vector, however, does not belong to $\mathrm{Im}_{\pm}G_3$.

To further constrain the structure of $\mathcal{C}$, we can require that its codewords are the solutions of an autoregressive system of equations, i.e. there exists a finite set of matrices $H_{ij} \in \mathbf{F}^{q\times n}$, such that $\mathbf{w} = \sum_{h,k\in\mathbf{Z}} w(h,k) z_1^h z_2^k$ belongs to $\mathcal{C}$ if and only if

$$\sum_{i,j} w(h-i, k-j) H_{ij}^T = 0, \qquad \forall\ (h,k) \in \mathbf{Z}\times\mathbf{Z}. \tag{2.3}$$

Thus, letting $H^T(z_1, z_2) := \sum_{i,j} H_{ij} z_1^i z_2^j = [h_1^T(z_1,z_2)\ \ldots\ h_q^T(z_1,z_2)]$, $\mathbf{w} \in \mathcal{C}$ if and only if

$$\mathbf{w} H^T(z_1, z_2) = 0. \tag{2.4}$$

Each column of H^T provides a *parity check*, which can be applied to a received sequence for testing whether it belongs to the code, and the representation

$$\mathcal{C} = \ker_{\pm} H^T := \{\mathbf{w} \in \mathcal{F}_{\pm}^n : \mathbf{w} H^T(z_1, z_2) = 0\}$$

shows that a finite number of parity checks is sufficient for a complete characterization of $\mathcal{C}$. The matrix $H^T(z_1, z_2)$ will be referred to as a *syndrome decoder* of $\mathcal{C}$, and the corresponding codes are called *finite convolutional codes*. Their structure is characterized by the following proposition.

Proposition 1 A free modular code $\mathcal{C}$ admits a syndrome decoder if and only if $\mathcal{C}$ has a ℓFP encoder $\bar{G}(z_1, z_2)$.

Proof Let $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G}$, where $\bar{G} \in \mathcal{F}_{\pm}^{k\times n}$ is ℓFP, and consider a full column rank matrix $H^T(z_2, z_2) \in \mathcal{F}_{\pm}^{n\times(n-k)}$, such that $\bar{G}H^T = 0$.
Clearly, if $\mathbf{w} \in \mathcal{C}$, then $\mathbf{w} = \mathbf{u}\bar{G}$, for some $\mathbf{u} \in \mathcal{F}_{\pm}^k$, and $\mathbf{w}H^T = (\mathbf{u}\bar{G})H^T = \mathbf{u}(\bar{G}H^T) = 0$, so $\mathbf{w} \in \ker_{\pm} H^T$. On the other hand, if $\mathbf{w} \in \mathcal{F}_{\pm}^k$ is in $\ker_{\pm} H^T$, it belongs to the subspace of $\mathbf{F}(z_1, z_2)^n$ orthogonal to the columns of H^T, and spanned by the rows of $\bar{G}$. Then there exists a row vector $\mathbf{f} \in \mathbf{F}(z_1, z_2)^k$ such that

$$\mathbf{w} = \mathbf{f}\bar{G}(z_1, z_2). \tag{2.5}$$

We aim to prove that $\mathbf{f}$ is an element of $\mathcal{F}_{\pm}^k$. Actually, as $\bar{G}$ is ℓFP, there exist [8] two L-polynomials $h(z_1) \in \mathbf{F}[z_1, z_1^{-1}]$, $k(z_2) \in \mathbf{F}[z_2, z_2^{-1}]$, and two L-polynomial matrices $X(z_1, z_2)$ and $Y(z_1, z_2)$, such that

$$\bar{G}(z_1, z_2)X(z_1, z_2) = h(z_1)I_k \quad \text{and} \quad \bar{G}(z_1, z_2)Y(z_1, z_2) = k(z_2)I_k. \tag{2.6}$$

It entails no loss of generality supposing that $\mathbf{f}$ has irreducible entries, f_i. So, letting $d(z_1, z_2)$ the l.c.m. of the denominators of f_i, (2.5) can be rewritten as

$$d\ \mathbf{w} = [n_1\ ...\ n_k]\ \bar{G}, \qquad n_i \in \mathcal{F}_{\pm}, \quad i = 1, 2, ... \tag{2.7}$$

Postmultiplying both members of (2.7) by $X(z_1, z_2)$ and $Y(z_1, z_2)$, we obtain

$$\begin{aligned} d\ \mathbf{w}X(z_1, z_2) &= [n_1\ ...\ n_k]\ \bar{G}X(z_1, z_2) = [n_1\ ...\ n_k]\ h(z_1) \\ d\ \mathbf{w}Y(z_1, z_2) &= [n_1\ ...\ n_k]\ \bar{G}Y(z_1, z_2) = [n_1\ ...\ n_k]\ k(z_2), \end{aligned}$$

respectively. As $d, n_1, ..., n_k$ have no common factors, it follows that $d(z_1, z_2) \mid h(z_1)$ and $d(z_1, z_2) \mid k(z_2)$, and therefore d is a unit in $\mathcal{F}_{\pm}$. Thus $\mathbf{f}$ belongs to $\mathcal{F}_{\pm}^k$ and $\mathbf{w}$ to $\mathrm{Im}_{\pm}\bar{G}$.

Vice versa, let $\mathcal{C} = \ker_{\pm} H^T$, with $H^T \in \mathcal{F}_{\pm}^{n\times p}$ and rank r, and consider any ℓFP $\bar{G}(z_1, z_2) \in \mathcal{F}_{\pm}^{(n-r)\times n}$, such that $\bar{G}H^T = 0$. Using the same arguments as in the first part of the proof, one shows that $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G}$ ■

Remark By the above proof, given any encoder G of a finite code $\mathcal{C}$, each set of generators for the subspace of $\mathbf{F}(z_1, z_2)^n$ orthogonal to the rows of G, constitutes a syndrome decoder of $\mathcal{C}$. In particular, we can always resort to a rFP syndrome decoder H^T, which is unique modulo a right unimodular factor.

Some specific reliability requirements, concerning the reconstruction of the information sequences at the decoding stage, justify the introduction of our restriction on the structure of $\mathcal{C}$. Usually, the received sequence $\mathbf{w}_r$ is not in $\mathcal{C}$ but, when the transmission system is well-designed, $\mathbf{w}_r$ differs from a codeword $\mathbf{w}$ of $\mathcal{C}$ in a finite number of points, and therefore the error sequence $\mathbf{e} := \mathbf{w}_r - \mathbf{w}$ belongs to $\mathcal{F}_{\pm}^n$. Since an injective encoder $G \in \mathcal{F}_{\pm}^{k\times n}$ induces a bijection between $\mathcal{F}_{\pm}^k$ and $\mathcal{C}$, there exists a decoder $G^{-1}(z_1, z_2) \in \mathbf{F}(z_1, z_2)^{n\times k}$ such that $GG^{-1} = I_k$. So, when restricted to the set of the codewords $\mathcal{C}$, $G^{-1}(z_1, z_2)$ represents the inverse of the encoding map. The error sequence $\mathbf{e}$, however, needs not to be a codeword so applying $G^{-1}(z_1, z_2)$ to $\mathbf{w}_r$ gives back the sequence

$$\mathbf{u}_r = \mathbf{w}_r G^{-1} = (\mathbf{u}G)G^{-1} + \mathbf{e}G^{-1} = \mathbf{u} + \mathbf{e}G^{-1},$$

which differs from the original information sequence by the (possibly infinite) reconstruction error $\mathbf{e}G^{-1} = \mathbf{u}_r - \mathbf{u}$.

To avoid this kind of catastrophic errors, it is imperative to use an L-polynomials decoder, that exists if and only if $\mathcal{C}$ admits a ℓZP encoder G. Analogously with the 1D case, modular codes generated by a ℓZP polynomial matrix will be called *finite basic codes*.

Example Let $\mathbf{F} = GF(2)$. It's easy to check that the following L-polynomial matrix

$$G_1(z_1, z_2) = \begin{bmatrix} z_1^{-1} + 1 & 0 & z_1^2 \\ z_2^{-1} & z_2 + 1 & 0 \end{bmatrix}$$

is ℓZP, while

$$G_2(z_1, z_2) = \begin{bmatrix} z_1^2 + 1 & 0 & z_1 \\ z_2 + 1 & z_2^2 + z_1 & 0 \end{bmatrix}$$

is ℓFP but not ℓZP, since its maximal order minors have a common zero in $(1, 1)$. Therefore, no L-polynomial right inverse of G_2 exists.

As 2D finite basic codes constitute a proper subclass of convolutional ones, it might be expected that a characterization of their structure should be possible also in terms of syndrome decoders. This is actually the case, as stated in the following proposition.

Proposition 2 Let $\mathcal{C}$ be a modular code. The followings are equivalent:

(i) $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G}$, with $\bar{G} \in \mathcal{F}_{\pm}^{k\times n}$ and ℓZP;

(ii) $\mathcal{C} = \ker\bar{H}^T$, with $\bar{H}^T \in \mathcal{F}_{\pm}^{n\times(n-k)}$ and rZP.

Proof $(i) \Rightarrow (ii)$ By the Quillen-Suslin theorem [11], there exists an L-polynomial matrix $\bar{P}(z_1, z_2)$ such that

$$U(z_1, z_2) := \begin{bmatrix} \bar{G}(z_1, z_2) \\ \bar{P}(z_1, z_2) \end{bmatrix}$$

is unimodular. The rZP matrix $\bar{H}^T(z_1, z_2) \in \mathcal{F}_{\pm}^{n\times(n-k)}$, constituted by the last $n - k$ columns of the inverse matrix $U^{-1}(z_1, z_2) = [\bar{L}^T(z_1, z_2) \quad \bar{H}^T(z_1, z_2)]$, satisfies $\bar{G}\bar{H}^T = 0$, and therefore is a syndrome decoder of $\mathcal{C}$.

$(ii) \Rightarrow (i)$ Using the same argument as in the first part of the proof, $\bar{H}^T(z_1, z_2)$ can be column-bordered into a unimodular matrix $V(z_1, z_2) := [\bar{L}^T(z_1, z_2) \quad \bar{H}^T(z_1, z_2)]$. The first k rows of $V^{-1}(z_1, z_2)$ provide a ℓZP encoder $\bar{G}(z_1, z_2)$ of $\mathcal{C}$ ■

3 Equivalent encoders

The above discussion made it clear that a modular code can be generated by different encoders. In a more algebraic theoretic setting, this amounts to say that an $\mathcal{F}_{\pm}$-module admits different families of generators.

Two matrices $G_1 \in \mathcal{F}_{\pm}^{k_1\times n}$ and $G_2 \in \mathcal{F}_{\pm}^{k_2\times n}$ are *equivalent encoders* ($G_1 \sim G_2$) if they generate the same code, i.e. if the $\mathcal{F}_{\pm}$- modules generated by the rows of G_1 and G_2 coincide. This implies that G_1 is equivalent to G_2 if and only if there exist two L-polynomial matrices $P_1 \in \mathcal{F}_{\pm}^{k_2\times k_1}$ and $P_2 \in \mathcal{F}_{\pm}^{k_1\times k_2}$ such that

$$P_1 G_1 = G_2 \qquad P_2 G_2 = G_1 \tag{3.1}$$

When confining ourselves to the class of full row rank encoders (namely, the injective encoders of free modular codes), we can replace (3.1) with the single equation

$$G_1 = U G_2, \tag{3.2}$$

where $U(z_1, z_2)$ denotes an $\mathcal{F}_\pm$-unimodular matrix. Indeed, (3.1) and the row rank assumption on G_1 and G_2 imply that both matrices have the same number, say k, of rows, and P_1 and P_2 are $k \times k$ L-polynomial matrices. From $G_1 = P_2 G_2 = P_2 P_1 G_1$ we get $P_2 P_1 = I_k$ and consequently $U(z_1, z_2) := P_2(z_1, z_2)$ is $\mathcal{F}_\pm$-unimodular. So, when a code $\mathcal{C}$ admits a ℓFP (ℓZP) encoder, all the injective encoders of $\mathcal{C}$ are ℓFP (ℓZP), too.

As the various subclasses of modular codes introduced in section 2 are characterized by the existence of suitable (injective, ℓFP or ℓZP) encoders, an important issue is to decide whether a code $\mathcal{C}$, given through the assignment of an arbitrary encoder G, admits an encoder enjoying the aforementioned rank and primeness properties. The following proposition provides a complete answer.

Proposition 3 Let $G(z_1, z_2)$ be in $\mathcal{F}_\pm^{k\times n}$, with rank $\bar{k}$ over $\mathbf{F}(z_1, z_2)$. Then there exist two L-polynomial matrices, $\bar{G}(z_1, z_2) \in \mathcal{F}_\pm^{\bar{k}\times n}$ ℓFP and $T(z_1, z_2) \in \mathcal{F}_\pm^{k\times \bar{k}}$ with full column rank, such that

$$G(z_1, z_2) = T(z_1, z_2)\bar{G}(z_1, z_2). \tag{3.3}$$

Moreover, the code $\mathcal{C} = \mathrm{Im}_\pm G$

(i) is free modular if and only if T factorizes into the product

$$T(z_1, z_2) = \bar{T}(z_1, z_2)L(z_1, z_2) \tag{3.4}$$

where $\bar{T}$ is rZP and L is a non singular square matrix;

(ii) is finite convolutional if and only if T is rZP;

(iii) is finite basic if and only if T is rZP and $\bar{G}$ is ℓZP.

Proof Let G' be a $\bar{k} \times n$ L-polynomial matrix, obtained by selecting in G $\bar{k}$ rows linearly independent over $\mathbf{F}(z_1, z_2)$. Then $G = RG'$, $R \in \mathbf{F}(z_1, z_2)^{k\times\bar{k}}$. Consider any g.l.f. Q of G' and factorize G' into $Q\bar{G}$, $\bar{G} \in \mathcal{F}_\pm^{\bar{k}\times n}$ ℓFP. So $G = T\bar{G}$, where $T = RQ$ is an L-polynomial matrix, by the same reasonings as in the proof of Proposition 1.

(i) Assume that in (3.4) $\bar{T}$ is rZP and L is a nonsingular square L-polynomial matrix, and consider the factorization $G = \bar{T}(L\bar{G})$. As $\bar{T}$ is right zero prime, the map $\bar{T} : \mathcal{F}_\pm^k \to \mathcal{F}_\pm^{\bar{k}}$ is surjective, and we have $\mathrm{Im}_\pm G = \mathrm{Im}_\pm L\bar{G}$. Being the image of a full row rank matrix, the code $\mathcal{C}$ is free modular.

Vice versa, let $\mathcal{C} = \mathrm{Im}_\pm G$ be a free $\mathcal{F}_\pm$-module. Then, there exist a full row rank L-polynomial matrix $\tilde{G}$ such that $\mathrm{Im}_\pm G = \mathrm{Im}_\pm \tilde{G}$, and two L-polynomial matrices P and $\tilde{P}$ satisfying

$$G = \tilde{P}\tilde{G}, \qquad \tilde{G} = PG. \tag{3.5}$$

From (3.5) one gets

$$(P\tilde{P} - I)\tilde{G} = 0, \tag{3.6}$$

and the row rank assumption on $\tilde{G}$ implies $P\tilde{P} = I$. So P is ℓZP and $\tilde{P}$ is rZP.

On the other hand, factorize the matrix T appearing in (3.3) as $T = \bar{T}L$, where $\bar{T}$ is rFP. Using (3.5), we get $\bar{T}(L\bar{G}) = G = \tilde{P}\tilde{G} = \tilde{P}PG = \tilde{P}P\bar{T}(L\bar{G})$, and consequently $\bar{T} = \tilde{P}[P\bar{T}]$. As T is rFP and $\tilde{P}$ is rZP, it follows that $P\bar{T}$ is unimodular and $\bar{T}$ is rZP.

(ii) and *(iii)* If in (3.3) T is a rZP matrix, the map $T : \mathcal{F}_{\pm}^{k} \to \mathcal{F}_{\pm}^{\bar{k}}$ is surjective and therefore $\mathrm{Im}_{\pm}G = \mathrm{Im}_{\pm}\bar{G}$. This means that $\mathcal{C}$ is finite convolutional when $\bar{G}$ is ℓFP, and finite basic when $\bar{G}$ is ℓZP.
Conversely, if $\mathrm{Im}_{\pm}G = \mathrm{Im}_{\pm}\tilde{G}$ for some ℓFP (ℓZP) $\bar{k} \times n$ matrix $\tilde{G}$, there exists an L-polynomial matrix P such that $\tilde{G} = PG$, and therefore

$$\tilde{G} = (PT)\bar{G}. \tag{3.7}$$

As both $\bar{G}$ and $\tilde{G}$ are ℓFP, PT is unimodular and T is rZP. Moreover, if $\tilde{G}$ is ℓZP, $\bar{G}$ is ℓZP too. ■

In the remaining part of this section, we shall confine ourselves to finite convolutional codes. Since these codes can be characterized as kernels of syndrome decoders, it seems quite natural to ask how two syndrome decoders of the same code $\mathcal{C}$ are related each other. The following proposition provides an equivalence condition for two syndrome decoders, and shows that, when dealing with encoders of convolutional codes, the equivalence condition (3.1) can be replaced by a single L-polynomial equation.

Proposition 4 Consider a pair of finite convolutional codes $\mathcal{C}_i = \mathrm{Im}_{\pm}G_i = \ker_{\pm}H_i^T$, $i = 1, 2$. Then $\mathcal{C}_1 = \mathcal{C}_2$ if and only if

a) there exist two full column rank L-polynomial matrices P_1 and P_2 such that

$$P_1G_1 = P_2G_2 \tag{3.8}$$

or, equivalently,

b) there exist two full row rank L-polynomial matrices Q_1 and Q_2 such that

$$H_1^TQ_1 = H_2^TQ_2 \tag{3.9}$$

Proof a) Assume first $\mathrm{Im}_{\pm}G_1 = \mathrm{Im}_{\pm}G_2$. By Proposition 3, there exist two rZP L-polynomial matrices T_1 and T_2 such that $G_i = T_i\bar{G}_i$, $\bar{G}_i$ ℓFP, i=1,2. *Sincewehave* $\mathrm{Im}_{\pm}\bar{G}_1 = \mathrm{Im}_{\pm}G_1 = \mathrm{Im}_{\pm}G_2 = \mathrm{Im}_{\pm}\bar{G}_2$, we can find an $\mathcal{F}_{\pm}$-unimodular matrix $U(z_1, z_2)$, satisfying $\bar{G}_1 = U\bar{G}_2$, which, in turn, gives

$$T_1^{-1}(T_1\bar{G}_1) = UT_2^{-1}(T_2\bar{G}_2), \tag{3.10}$$

T_1^{-1} and T_2^{-1} L-polynomial left inverses of T_1 and T_2, respectively. Putting $P_1 := T_1^{-1}$ and $P_2 := UT_2^{-1}$ in (3.10), one gets equation (3.8).

Viceversa, assume that (3.8) holds and, using Proposition 3, let $G_i = T_i\bar{G}_i$, T_i rZP, $\bar{G}_i$ ℓFP, $i = 1, 2$. This gives $(P_1T_1)\bar{G}_1 = (P_2T_2)\bar{G}_2$, and, consequently,

$$\bar{G}_1 = (T_1^{-1}P_1^{-1}P_2T_2)\bar{G}_2, \tag{3.11}$$

where T_1^{-1} and P_1^{-1} are rational left inverses of T_1 and P_1 respectively. As both $\bar{G}_1$ and $\bar{G}_2$ are ℓFP, $T_1^{-1}P_1^{-1}P_2T_2$ is an $\mathcal{F}_\pm$-unimodular matrix. So, the equivalence chain $G_1 \sim \bar{G}_1 \sim \bar{G}_2 \sim G_2$ proves that G_1 and G_2 are equivalent encoders.

b) If (3.9) holds, we have

$$\mathbf{w}H_1^TQ_1 = 0 \;\Leftrightarrow\; \mathbf{w}H_1^T = 0$$

and, similarly,

$$\mathbf{w}H_2^TQ_2 = 0 \;\Leftrightarrow\; \mathbf{w}H_2^T = 0.$$

Therefore

$$\ker_\pm H_1^T = \ker_\pm H_1^TQ_1 = \ker_\pm H_2^TQ_2 = \ker_\pm H_2^T.$$

Viceversa, if H_1^T and H_2^T are equivalent syndrome decoders, the columns of H_1^T and H_2^T generate the same subspace in $\mathbf{F}^n(z_1, z_2)$. Hence, there exists a rational matrix L that satisfies the equation $H_1^TL = H_2^T$. We can column-border L into a full row rank matrix $[L \;\; M]$, so as to get

$$H_1^T\,[\,L \quad M\,] = H_2^T\,[\,I \quad N\,]\,, \tag{3.12}$$

where M and N are suitable rational matrices. Consider now any rMFD RS^{-1} of $[\,L \quad M\,]$, and rewrite (3.12) as $H_1^TR = H_2^T\,[\,I \quad N\,]\,S$. R is clearly full row rank and, denoting by Q_2J^{-1} any rMFD of $[\,I \quad N\,]\,S$, we get $H_1^TQ_1 = H_1^TRJ = H_2^TQ_2$, where both $Q_1 := RJ$ and Q_2 are full row rank ■

4 Dual codes

An obvious way to extend the finite codes considered in the previous sections, is to relax the constraints on the supports of the codewords, thus allowing the codes to include sequences with infinite supports. This point of view has been adopted in [6], where (infinite) convolutional codes have been introduced by imposing increasingly stronger constraints, typical of the "behavioural approach" [3,9], on two-dimensional sequences in $\mathcal{F}_\infty^n$.

In this section we aim to show that every complete and, in particular, convolutional (infinite) code can be seen as the set of all parity checks that can be applied to an arbitrary sequence of $\mathcal{F}_\pm^n$, to decide whether it belongs to a given modular code $\mathcal{C}$.
From an algebraic point of view, this amounts to regard an infinite code as a space of linear functionals on $\mathcal{F}_\pm^n$, i.e. as the algebraic dual of a modular code.

Introduce in $\mathcal{F}_\pm^m \times \mathcal{F}_\infty^m$ the non degenerate bilinear form

$$\langle\cdot,\cdot\rangle_m : \mathcal{F}_\pm^m \times \mathcal{F}_\infty^m \to \mathbf{F},$$

defined by : $\langle\mathbf{u},\mathbf{v}\rangle_m = (\mathbf{u}\mathbf{v}^T, 1) = \sum_{i,j\in\mathbf{Z}} u(i,j)v^T(-i,-j)$.
Two vectors $\mathbf{u} \in \mathcal{F}_\pm^m$ and $\mathbf{v} \in \mathcal{F}_\infty^m$ are called *orthogonal* if $\langle\mathbf{u},\mathbf{v}\rangle_m = 0$. Given any modular code $\mathcal{C} \subseteq \mathcal{F}_\pm^m$, its orthogonal complement $\mathcal{C}^\perp$ is constituted by all vectors

of $\mathcal{F}_\pm^m$ which are orthogonal to $\mathcal{C}$. Similarly, every submodule $\mathcal{D}$ of $\mathcal{F}_\infty^m$ identifies an orthogonal complement $\mathcal{D}^\perp$ in $\mathcal{F}_\pm^m$.
We can associate with every $\mathbf{v} \in \mathcal{F}_\pm^m$ the linear functional on $\mathcal{F}_\pm^m$ defined by

$$f_v(\cdot) = \langle \cdot, \mathbf{v} \rangle_m \tag{4.1}$$

and, conversely, every linear functional on $\mathcal{F}_\pm^m$ can be represented as in (4.1), for an appropriate choice of $\mathbf{v} \in \mathcal{F}_\infty^m$. This way the space $\mathcal{F}_\infty^m$ is identified with $L(\mathcal{F}_\pm^m)$, and several strong results, which do not hold for arbitrary pairs of dual spaces, are made available [10].

Let $\mathcal{C}$ be a modular code, described as the image of the map

$$G : \mathcal{F}_\pm^k \to \mathcal{F}_\pm^n : \mathbf{u} \mapsto \mathbf{u}G,$$

and consider the map

$$G^T : \mathcal{F}_\infty^n \to \mathcal{F}_\infty^k : \mathbf{v} \mapsto \mathbf{v}G^T.$$

As $\langle \mathbf{u}G, \mathbf{v} \rangle_n = (\mathbf{u}G\mathbf{v}^T, 1) = (\mathbf{u}(\mathbf{v}G^T)^T, 1) = \langle \mathbf{u}, \mathbf{v}G^T \rangle_k$, then G and G^T are dual mappings. This implies

$$\mathrm{Im}_\pm G = (\mathrm{ker}G^T)^\perp, \tag{4.2}$$

and

$$\mathrm{ker}G^T = (\mathrm{Im}_\pm G)^\perp, \tag{4.3}$$

where

$$\mathrm{ker}G^T := \{\mathbf{v} \in \mathcal{F}_\infty^n : \mathbf{v}G^T = 0\}. \tag{4.4}$$

By (4.2), the $\mathcal{F}_\pm$-submodule of $\mathcal{F}_\infty^n$, $\mathcal{D} := \mathrm{ker}G^T$, represents the set of all linear functions $f_\mathbf{v}(\cdot)$ we are allowed to apply when deciding whether $\mathbf{w} \in \mathcal{F}_\infty^n$ belongs to $\mathcal{C}$, and it will be called the *dual code* of $\mathcal{C}$.
Relations (4.2) and (4.3) together, induce a bijective map between the family of modular codes (i.e. the family of submodules of $\mathcal{F}_\pm^n$) and a family of specific $\mathcal{F}_\pm$-submodules of $\mathcal{F}_\infty^n$, namely those that can be described as the kernel of polynomial operators. In the sequel we will analyse some "internal" properties which characterize infinite codes that can be described as duals of modular codes. Moreover we aim to investigate how the subclasses of modular codes considered in section 2, mirror into classes of dual codes having very special structures.

The submodules of $\mathcal{F}_\infty^n$ which can be represented as the kernel of a polynomial matrix, are exactly those which are close in the pointwise convergence topology, i.e. the so called "complete dual codes" [6]. A complete dual code $\mathcal{D}$ can be characterized as follows: given an infinite sequence $\mathcal{S}_1 \subset \mathcal{S}_2 \subset ...$ of finite windows invading $\mathbf{Z} \times \mathbf{Z}$ (so that every point $(i, j) \in \mathbf{Z} \times \mathbf{Z}$ eventually belongs to all the windows of the sequence), a sequence $\mathbf{v} \in \mathcal{F}_\infty^n$ is an element of $\mathcal{D}$ if and only if there exist codewords $\mathbf{v}_1, \mathbf{v}_2, ...$ in $\mathcal{D}$ such that $\mathbf{v}_i|\mathcal{S}_i = \mathbf{v}|\mathcal{S}_i, \quad i = 1, 2, ...$
In general, to test whether $\mathbf{w} \in \mathcal{F}_\pm^n$ is in some modular code $\mathcal{C}$, we resort to parity checks represented by elements of $\mathcal{D}$, which possibly have an infinite support. This kind

of checks seem quite unsuitable for an algorithmic implementation, so it's interesting to determine when the submodule of the finite codewords of the dual code,

$$\mathcal{D}_f := \{\mathbf{v} \in \mathcal{D} : \mathbf{v} \in \mathcal{F}_\pm^n\} = \mathcal{D} \cap \mathcal{F}_\pm^n \tag{4.5}$$

constitutes a set of parity checks sufficient to decide whether $\mathbf{w}$ is in $\mathcal{C}$, namely under which conditions the equivalence

$$\mathbf{w} \in \mathcal{C} \qquad \Leftrightarrow \qquad \langle \mathbf{w}, \mathbf{v} \rangle_n = 0, \qquad \forall\ \mathbf{v} \in \mathcal{D}_f \tag{4.6}$$

holds. Being an $\mathcal{F}_\pm$-submodule of $\mathcal{F}_\pm^n$, $\mathcal{D}_f$ is finitely generated, that is $\mathcal{D}_f = \mathrm{Im}_\pm H$ for some $p \times n$ L-polynomial matrix. Thus (4.6) can be restated as

$$\mathbf{w} \in \mathcal{C} \qquad \Leftrightarrow \qquad \mathbf{w} H^T = 0, \tag{4.7}$$

where $H^T(z_1, z_2)$ can be seen as a syndrome decoder. As shown in section 2, a syndrome decoder of $\mathcal{C}$ can be found if and only $\mathcal{C}$ is the image of a ℓFP L-polynomial matrix. Therefore the submodule $\mathcal{D}_f$ of the dual code $\mathcal{D} = \mathcal{C}^\perp$ provides a set of parity checks, rich enough to identify the elements of $\mathcal{C}$ if and only if $\mathcal{C}$ is finite convolutional, in this case it's natural to expect that the whole dual code $\mathcal{D}$ can be uniquely reconstructed from $\mathcal{D}_f$. The following proposition shows that this is true, indeed, and analyses how the main features of finite convolutional codes translate, via duality, into properties of the corresponding duals, that will be called *dual convolutional codes*.

Proposition 5 [Finite and dual convolutional codes] Let $\mathcal{C}$ be a modular code of length n and $\mathcal{D} = \mathcal{C}^\perp$ its dual. The following facts are equivalent:

(a) $\mathcal{C}$ is finite convolutional, i.e. $\mathcal{C} = \ker_\pm H^T$ for some L-polynomial matrix H^T;

(b) $\mathcal{C} = \mathrm{Im}_\pm \bar{G}$, for some $\bar{G}$ ℓFP L-polynomial matrix;

(c) $\mathcal{D} = \mathrm{Im}\tilde{H} := \{\mathbf{v} \in \mathcal{F}_\infty^n : \mathbf{v} = \mathbf{u}\tilde{H}, \mathbf{u} \in \mathcal{F}_\infty^p\}$ for some L-polynomial matrix $\tilde{H}$;

(d) $\mathcal{D} = \ker\tilde{G}^T = \{\mathbf{v} \in \mathcal{F}_\infty^n : \mathbf{v}\tilde{G}^T = 0\}$ for some $\tilde{G}^T$ ℓFP;

(e) $\mathcal{D}$ is the closure, in the pointwise convergence topology on $\mathcal{F}_\infty^n$, of the $\mathcal{F}_\pm$-module $\mathcal{D}_f$.

Proof (a) $\Leftrightarrow$ (b) has been proved in section 2.
By resorting to the well-known property of dual maps

$$(\ker_\pm H^T)^\perp = \mathrm{Im} H, \tag{4.8}$$

one gets $\mathcal{D} = \mathcal{C}^\perp = (\ker_\pm H^T)^\perp = \mathrm{Im}H$, so that (a) $\Rightarrow$ (c), while

$$(\mathrm{Im}\tilde{H})^\perp = \ker_\pm \tilde{H}^T \tag{4.9}$$

implies $\mathcal{C} = \mathcal{D}^\perp = (\mathrm{Im}\tilde{H})^\perp = \ker_\pm \tilde{H}^T$, and hence (c) $\Rightarrow$ (a).

Analogously, from (4.3) it follows that $\mathcal{D} = \mathcal{C}^{\perp} = (\mathrm{Im}_{\pm}\bar{G})^{\perp} = \ker\bar{G}^T$ and therefore (b) $\Rightarrow$ (d), whereas, from (4.2) one gets $\mathcal{C} = \mathcal{D}^{\perp} = (\ker\tilde{\bar{G}}^T)^{\perp} = \mathrm{Im}_{\pm}\tilde{\bar{G}}$, and so (d) $\Rightarrow$ (b). Finally, the equivalence (c) $\Leftrightarrow$ (e) has been proved in [6,9]. ■

Remark If $\mathcal{C} = \mathrm{Im}_{\pm}G$ is not a finite convolutional code, i.e. $G = T\bar{G}$, with $\bar{G}(z_1, z_2)$ ℓFP and $T(z_1, z_2)$ a full column rank L-polynomial matrix, which is not rZP, by applying to a finite sequence $\mathbf{w}$ the parity checks associated with the elements of $\mathcal{D}_f$, we cannot guarantee that $\mathbf{w}$ is in $\mathcal{C}$. Indeed, the elements of $\mathcal{F}_{\pm}^n$ which belong to $\ker_{\pm}H^T$ are exactly the codewords of $\mathrm{Im}_{\pm}\bar{G}$.
Actually, as the rows of G belong to $\mathcal{C}$, then $0 = GH^T = T\bar{G}H^T$. Since T is a full column rank matrix, it follows that $\bar{G}(z_1, z_2)H^T(z_1, z_2) = 0$, and therefore $\mathrm{Im}_{\pm}\bar{G} \subseteq \ker_{\pm}H^T$.
Conversely, as the columns of H^T span in $\mathbf{F}^n(z_1, z_2)$ the vector space orthogonal to the rows of $\bar{G}$, each vector $\mathbf{w} \in \ker_{\pm}H^T$ can be expressed as a linear combination over $\mathbf{F}(z_1, z_2)$ of the rows of $\bar{G}$. By the left factor primeness of $\bar{G}$, the coefficients of the combination are in $\mathcal{F}_{\pm}$, namely $\mathbf{w} \in \mathrm{Im}_{\pm}\bar{G}$.
It's worthwhile to remark that the code $\mathrm{Im}_{\pm}\bar{G}$ is the minimal finite convolutional code including $\mathcal{C}$. Actually, if $\tilde{G}(z_1, z_2)$ is a ℓFP L-polynomial matrix such that $\mathcal{C} \subseteq \mathrm{Im}_{\pm}\tilde{G}$, there exists an L-polynomial matrix $P(z_1, z_2)$ such that $P\tilde{G} = G = T\bar{G}$. As T is full column rank, there exists a left rational inverse $T^{-1}(z_1, z_2)$, so that $\bar{G} = (T^{-1}P)\tilde{G}$. Moreover, since $\tilde{G}$ is ℓFP and $\bar{G}$ L-polynomial, $T^{-1}P$ is an L-polynomial matrix, which implies that $\mathrm{Im}_{\pm}\bar{G} \subseteq \mathrm{Im}_{\pm}\tilde{G}$.

Example Let $\mathbf{F} = GF(2)$ and let $\mathcal{C} = \mathrm{Im}_{\pm}G$, where

$$G(z_1, z_2) = \begin{bmatrix} z_1 & z_2 + 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & z_2 & 1 \\ z_1 & 1 & 0 \end{bmatrix} =: T(z_1, z_2)\bar{G}(z_1, z_2).$$

A basis for the space orthogonal in $\mathbf{F}^3(z_1, z_2)$ to the rows of G, consists of the following vector

$$H(z_1, z_2) := [1 + z_2(1 + z_2 + z_1z_2) \quad z_1 + z_2 \quad 1 + z_2 + z_1z_2 + z_1z_2^2].$$

H^T is a syndrome decoder for the code $\bar{\mathcal{C}} := \mathrm{Im}_{\pm}\bar{G}$, which properly includes $\mathcal{C}$. Actually $\bar{\mathcal{C}}$ includes the sequence $\mathbf{w} = [z_1 \;\; 1 \;\; 0]$, which is not an element of $\mathcal{C}$, but produces an all-zero pattern when applied to the syndrome decoder H^T. Therefore $\mathbf{w}$ is recognized by H^T as a codeword.
The reason why the syndrome decoder proves to be unreliable for identifying the elements of $\mathcal{C}$, is that the totality of the parity checks in $\mathrm{Im}H$ is a proper subset of the dual code $\mathcal{D} = \ker G^T$. For instance, the infinite sequence

$$\mathbf{v} := [\sum_{i \in \mathbf{Z}} z_2^i \;\; 0 \;\; \sum_{i \in \mathbf{Z}} z_2^i]$$

is an element of $\mathcal{D}$ which is not in $\mathrm{Im}H$.
Note that, by applying $\mathbf{v}$ to the sequence $\mathbf{w}$, we recognize it as an illegal sequence, since $f_{\bar{v}}(\cdot) = \langle \mathbf{w}, \bar{\mathbf{v}} \rangle_n \neq 0$, for each $\bar{\mathbf{v}} = z_1^n z_2^m \mathbf{v}$, $n, m \in \mathbf{Z}$.

Proposition 2, together with the dual relations (4.2) $\div$ (4.3) and (4.8) $\div$ (4.9), allows to obtain a characterization of finite basic codes and their duals, which is very close to that provided by Proposition 5 for convolutional codes.

Proposition 6 [Finite and dual basic codes] Let $\mathcal{C}$ be a modular code of length n and $\mathcal{D} = \mathcal{C}^{\perp}$ the corresponding dual code. The followings are equivalent

(a) $\mathcal{C}$ is a finite basic code, namely $\mathcal{C} = \mathrm{Im}_{\pm}\bar{G},\ \bar{G}$ ℓZP;

(b) $\mathcal{C} = \ker_{\pm}\bar{H}^T,\ \bar{H}^T$ rZP;

(c) $\mathcal{D} = \ker\tilde{G}^T,\ \tilde{G}^T$ rZP;

(d) $\mathcal{D} = \mathrm{Im}\tilde{H},\ \tilde{H}$ ℓZP ■

As underlined by Propositions 5 and 6, the bijective correspondence between modular codes and dual complete codes, maps, in particular, finite convolutional codes into dual convolutional codes. Consequently, internal properties of the different classes of modular codes mirror into internal properties of the corresponding classes of dual codes. The analysis of these properties has been carried out in [6] mainly for dual codes, while an internal characterization of the different classes of modular codes is still unavailable. Indeed, what seems interesting to understand, is what kind of mutual relations among codewords characterize a modular code, without taking into account the input-output map which underlies their generation.

5 References

1. G.D.Forney *Convolutional codes I: algebraic structure*, IEEE Trans. Inf.Th., vol. 16, n. 6, pp.720-738, 1970

2. G.D.Forney *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. of Control, vol.13, n.3, pp.493-521, 1975

3. J.C.Willems *Models for dynamics*, Dynamics Reported, vol.2, U.Kirchgaber and H.O.Walther eds., Wiley and Teubner, pp.171-269,1989

4. G.D.Forney, M.D.Trott *The dynamics of group codes: state spaces, trellis diagrams and canonical encoders*, to appear in IEEE Trans.Inf. Th.

5. H.A.Loeliger, G.D. Forney, T.Mittelholzer, M.D.Trott *Minimality and observability of group systems*, submitted, 1993

6. E.Fornasini, M.E.Valcher *Algebraic aspects of 2D convolutional codes*, submitted to IEEE Trans.Inf.Th., 1993

7. S.Lang *Algebra*, Addison-Wesley Publ.Comp., 1967

8. M.Morf, B.C.Lévy, S.Y.Kung *New results in 2D systems theory: part I*, Proc.of the IEEE, vol.65, pp.861-872, 1977

9. P.Rocha *Structure and representation of 2D systems*, Ph.D.Thesis, Rijksuniversiteit, Groningen, 1990

10. W.Greub *Linear algebra*, Springer-Verlag, 1975

11. D.C.Youla, P.F.Pickel *The Quillen-Suslin theorem*, IEEE Trans. Circ. and Sys., vol.31, pp.513-518, 1984