ELSEVIER

Discrete Applied Mathematics 111 (2001) 139–155

DISCRETE APPLIED MATHEMATICS

A convolutional-like approach to p-adic codes

N. Lagorce

Laboratoire d'Arithmétique, Calcul formel et Optimisation, Université de Limoges,
13243 rue Albert Thomas, 87060 Limoges Cedex, France

Received 25 June 1999; revised 9 February 2000; accepted 13 July 2000

Abstract

In this paper, we will see that codes defined over p-adic fields can be used in the same way as convolutional codes. We will prove some theoretical results concerning their encoders and show that, in practice, they can be encoded and decoded efficiently. © 2001 Published by Elsevier Science B.V.

Keywords: Error correcting codes; Convolutional codes; p-adic codes

0. Introduction

Convolutional codes are massively used in practical error-correcting devices. One reason for this is that algorithms which encode and decode them are efficient and easy to implement. The difference between block codes and convolutional codes is that, in a block code, the codeword transmitted at time t depends only on the message at time t whereas, for a convolutional code, the word transmitted at time t depends on a certain number of messages emitted during the preceding time indices. More precisely, a block code is defined over a finite field $\mathbb{F}$ and a convolutional code is defined over the field $\mathbb{F}((D))$ of Laurent series over $\mathbb{F}$. The role of the indeterminate D is to remember the past messages. The main justification of this paper is that we can do the same thing with other kinds of indeterminate, i.e. we can define codes similar to convolutional codes over every complete discrete valuation ring having a finite residue field, and notably over extension of p-adic fields. We will call these codes *CDVR-codes*.

Note that p-adic codes were first used in [10,1], but with a different meaning. Indeed, the main idea was that, if we have a block code over the ring $\mathbb{Z}_p$ of p-adic integers then, for every n, it induces a block code over the ring $\mathbb{Z}/(p^n)$, and that some properties of the parent code can be found again in its child. However, in the rest of this paper, we will only use the term "p-adic code" to mean "CDVR-code over $\mathbb{Q}_p$".

This paper is organized as follows. Section 1 is a small reminder concerning discrete valuation rings whose results are used to give, in Section 2, a definition of

PII: S0166-218X(00)00349-8

CDVR-codes. In Section 3, we will study more in detail CDVR-codes defined over the field $\mathbb{Q}_2$ of 2-adic numbers. We will see there that, for a given code, we can choose some encoders with good properties, and that we can encode and decode efficiently. In Section 4 we will briefly see what happens if we use a ramified extension of $\mathbb{Q}_2$. Section 5 is an enumeration of some properties of 2-adic codes.

1. Discrete valuation rings

This section is devoted to a brief study of complete discrete valuation rings whose residue field is finite. The most important result for the rest of this paper is the classification of these rings (Theorem 3).

Definition 1. An integral domain A is called a *discrete valuation ring* if it is a principal ideal domain having a unique maximal ideal.

Let A be a discrete valuation ring. If M is its maximal ideal then the quotient field $\mathbb{F}=A/M$ is called the *residue field* of A. As M is a principal ideal it has a generator element, say π, called an *uniformizer* (or *local parameter*) of A. Then every element x in $A-\{0\}$ can be written as $x=\pi^r u$ where $r\geqslant 0$ and u is a unit in A. The integer r is called the *valuation* of x, denoted by $v_A(x)$ and is, in fact, the greatest integer n such that x is an element of M^n. By extension, let $v_A(0)=+\infty$. Therefore the application v_A from A to $\mathbb{N}\cup\{+\infty\}$ is a valuation over A. This valuation is extended over K, the field of fractions of A, in the following way: if x and y are two elements of A, $y\neq 0$, let $v_K(x/y)=v_A(x)-v_A(y)$ be the valuation of x/y.

Moreover, if $a>1$ is a real number, the function $x\mapsto |x|$ defined by $|0|=0$ and $|x|=a^{-v_K(x)}$ is an absolute value on K. This absolute value induces a distance over K defined by

$$d_K: K\times K\to\mathbb{R}^+$$

$$(x,y)\mapsto d_K(x,y)=|x-y|.$$

We recall that if (E,d) is a metric space, then a sequence (a_n) of elements of E is a *Cauchy sequence* if, for all $\varepsilon>0$, there exists an integer N_ε, such that for $n>N_\varepsilon$ and $m>N_\varepsilon$, $d(a_n,a_m)<\varepsilon$.

Definition 2. A discrete valuation ring A is said *complete* if every Cauchy sequence of elements of A converges (for the topology induced by d_K) in A.

Let A be a complete discrete valuation ring, M its maximal ideal, $\mathbb{F}$ its residue field and K its field of fractions. Let S be a set of representatives of the cosets of $\mathbb{F}$ in A. Then

$$A=\left\{\sum_{i\geqslant 0} x_i\pi^i \mid x_i\in S\right\},$$

$$M = \left\{ \sum_{i>0} x_i \pi^i \mid x_i \in S \right\},$$

$$K = \left\{ \sum_{i \geqslant d} x_i \pi^i \mid x_i \in S,\ d \in \mathbb{Z} \right\}.$$

The following theorem completely classifies the complete discrete valuation rings having a finite residue field.

Theorem 3. *Let $\mathbb{F}$ be a finite field of characteristic p and A a complete discrete valuation ring having residue field $\mathbb{F}$. Only two cases are possible*:

(1) *If A has characteristic p then A is isomorphic to the ring $\mathbb{F}[[T]]$ of formal power series over $\mathbb{F}$. Its field of fractions is isomorphic to the field $\mathbb{F}((T))$ of Laurent series over $\mathbb{F}$.*

(2) *If A has characteristic zero then A is isomorphic to an extension of the ring $\mathbb{Z}_p$ of p-adic integers. Its field of fractions is isomorphic to an extension of the field $\mathbb{Q}_p$ of p-adic numbers.*

For a proof of this theorem see, for example, [9] or [8].

For an element $x = \sum_{i \geqslant r} x_i \pi^i$ of K, the *weight* of x, denoted by $w(x)$, is the number of its non-zero coefficients, that is $w(x) = \#\{i \in \mathbb{Z} \mid x_i \neq 0\}$.

2. CDVR-codes

Let $\mathbb{F}$ be a finite field of characteristic p. Let A be a complete discrete valuation ring (CDVR) whose residue field is $\mathbb{F}$ and K its field of fractions. We denote by A_f the subsemiring of A containing the series of A whose general term is ultimately null (i.e. finite weight series).

Definition 4. An (n,k) CDVR-code over K is a k-dimensional subspace of the vector space K^n having a basis consisting entirely of vectors from A_f^n.

If $\mathscr{C}$ is an (n,k) CDVR-code over K, an *encoder* of $\mathscr{C}$ is a $k \times n$ matrix over K whose rows form a basis of $\mathscr{C}$. This encoder is said *finite* if its coefficients are in A_f.

Let G be a finite encoder for $\mathscr{C}$. Let u be an element of K^k. The *codeword* corresponding to u in $\mathscr{C}$ is the element $x = uG$ in K^n.

According to the classification of field of fractions of complete discrete valuation rings (Theorem 3), two types of codes are enclosed in Definition 4. If K has characteristic p, then we find again the classical definition of convolutional codes as can be seen in, for example, [6]. If K has characteristic zero, then this definition identifies a new class of codes with coefficients over extensions of p-adic fields.

Although some properties can be obtained directly, without distinguishing between these two cases (notably a notion of “state space”), we will not talk about that here. However, in the next section, we focus on 2-adic codes and we will see some basic properties of their encoders. (Note that the choice of 2-adic was driven by practical constraints and that theoretical results can be generalized to p-adic codes in a straightforward way.)

3. 2-adic codes

Definition 5. A 2-adic code with parameters (n,k) is a CDVR-code over $\mathbb{Q}_2$, that is a k-dimensional subspace of $\mathbb{Q}_2^n$, having a basis consisting entirely of vectors from $\mathbb{N}^n$.

Note that, in the context of 2-adic codes, an encoder is said to be finite if all its components are elements of $\mathbb{N}$.

Clearly, a given 2-adic code $\mathscr{C}$ possess an infinite number of encoders. Some of them have good properties and others are to be eliminated absolutely. We will see this in the next subsection.

3.1. Finite encoders

In this section, we denote the set $\{x/2^t \mid x \in \mathbb{N},\ t \geqslant 0\}$ by $\hat{\mathbb{N}}$ and the set $\{x/2^t \mid x \in \mathbb{Z},\ t \geqslant 0\}$ by $\hat{\mathbb{Z}}$. Let $\mathscr{C}$ be an (n,k) 2-adic code.

In the encoding process, we try to avoid the case where the image of a finite weight word is an infinite weight word as, in this case, a finite number of errors during the transmission can cause an infinite number of errors after the decoding stage. Only the elements of $\hat{\mathbb{N}}$ have a finite weight and there exists two kinds of infinite weight words in $\mathbb{Q}_2$, the elements of $\hat{\mathbb{Z}} - \hat{\mathbb{N}}$ and the elements of $\mathbb{Q}_2 - \hat{\mathbb{Z}}$. (Note that negative elements of $\mathbb{Z}$ have an infinite weight as $-1 = \sum_{i \geqslant 0} 2^i$.) Conditions on the encoders can be given to eliminate these two types of series (Theorem 6 and Definition 7, respectively). Moreover, Definition 13 gives an even stronger property insuring that, for every word, there is no any time delay between encoding and effective transmission.

Theorem 6. *Let G be a finite encoder for $\mathscr{C}$. The following three conditions are equivalent*:

(1) *If $x = uG$ is an element of $\hat{\mathbb{N}}^n$ then u is necessarily in $\hat{\mathbb{Z}}^k$*,

(2) *The gcd of the kth-order minors* (*determinants of $k \times k$ sub-matrices*) *of G is a power of* 2,

(3) *G has a right inverse with coefficients in $\hat{\mathbb{Z}}$.*

An encoder satisfying these conditions is said to be weakly non-catastrophic.

Proof. We will prove successively $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$.

(1) ⇒ (2): We will, in fact, prove that if the gcd of the kth-order minors of G is not a power of 2, then there exists an element u not belonging to $\hat{\mathbb{Z}}^k$ such that $x = uG$ is an element of $\hat{\mathbb{N}}^n$.

Let p be an odd prime dividing the gcd of the kth-order minors of G. Let $\tilde{G}$ be the reduction of G modulo p. As every kth-order minor of G is null modulo p, the rank of $\tilde{G}$ is strictly less than k. So there exists a non-null vector $\tilde{U}$ such that $\tilde{U}\tilde{G} = 0$. Let U be the canonical injection of $\tilde{U}$ in $\mathbb{N}$. Therefore, every coefficient of the vector UG is a multiple of p. As at least one coefficient of U is not divisible by p, then the vector $u = 1/pU$ is not a vector of $\hat{\mathbb{Z}}^k$. Moreover, by construction of u, the vector $x = uG$ has all its coefficients in $\hat{\mathbb{N}}$.

(2) ⇒ (3): For $v = 1, \ldots, \binom{n}{k}$, denote by G_v the $k \times k$ sub-matrices of G and by Γ_v the corresponding minors. For $v = 1, \ldots, \binom{n}{k}$, let K_v be the adjoint matrix of G_v, such that $G_v K_v = \Gamma_v I_k$. Inserting $n - k$ null rows in an appropriate way in K_v, we obtain a matrix K'_v such that $GK'_v = \Gamma_v I_k$, and this for all v.

As the gcd of the Γ_v is of the form 2^l, for some $l \geqslant 0$, then there exists some integers h_v such that $\sum_v h_v \Gamma_v = 2^l$. Let

$$K' = \sum_v h_v K'_v.$$

Then every coefficient of K' belongs to $\mathbb{Z}$ and

$$\begin{aligned} GK' &= \sum_v h_v G K'_v \\ &= \sum_v h_v \Gamma_v I_k \\ &= \left(\sum_v h_v \Gamma_v \right) I_k \\ &= 2^l I_k. \end{aligned}$$

Therefore, the matrix $K = 2^{-l} K'$ is a right inverse of G having its coefficients in $\hat{\mathbb{Z}}$.

(3) ⇒ (1): Let H be a right inverse of G in $\hat{\mathbb{Z}}$. Suppose that $x = uG$ is an element of $\hat{\mathbb{N}}^n$. Then we have $u = xH$ and, as x has its coefficients over $\hat{\mathbb{N}}$ and H has its coefficients over $\hat{\mathbb{Z}}$, u is an element of $\hat{\mathbb{Z}}^k$. □

Definition 7. Let G be a weakly non-catastrophic encoder. G is said to be (*strongly*) *non-catastrophic* if, when $x = uG$ is an element of $\hat{\mathbb{N}}^n$, then necessarily u is an element of $\hat{\mathbb{N}}^k$.

Definition 8. Let G be a $k \times n$ encoder and $1 \leqslant i \leqslant k$. We say that G has an *information position* for i if there exists an index $1 \leqslant j \leqslant n$ such that, $g_{i,j} \neq 0$ and, for all $l \neq i$, $g_{l,j} = 0$.

Definition 9. A $k \times n$ encoder G is said to be *quasi-monomial* if it has an information position for every $i = 1,\dots,k$. In other words, G is a quasi-monomial encoder if and only if there exists a diagonal $k \times k$ matrix A, a $k \times (n-k)$ matrix B and a permutation matrix P of rank n such that

$$G = (A \mid B)P.$$

Theorem 10. *A weakly non-catastrophic encoder is strongly non-catastrophic if and only if it is quasi-monomial.*

For the proof of this theorem, we need the following two lemmas.

Lemma 11. *Let x and y be two vectors in $\hat{\mathbb{N}}^n$. There exists a non-negative integer α such that $\alpha x - y \geqslant 0$ if and only if, for all $i = 1,\dots,n$, $x_i = 0$ implies $y_i = 0$.*

Proof of Lemma 11

- Let us show that if $x_i = 0$ implies $y_i = 0$, for all i, then there exists an integer α such that $\alpha x - y \geqslant 0$. Let $1 \leqslant i \leqslant n$. Suppose that x_i is a non-zero element of $\hat{\mathbb{N}}$. Then there exists an integer s such that $2^s x_i$ and $2^s y_i$ are non-negative integers. According to Euclidean division in $\mathbb{Z}$, there exists $a \in \mathbb{N}$ and $b \in \mathbb{N}$ such that $2^s y_i = a2^s x_i + b$, with $0 \leqslant b < 2^s x_i$. Let $\alpha_i = a + 1$. Then $\alpha_i x_i \geqslant y_i$.
  If x_i is zero, then by hypothesis, y_i is also zero and $\alpha_i = 0$.
  Let $\alpha = \max_i \alpha_i$. Then $\alpha x - y \geqslant 0$.
- Let us suppose that there exists an integer α such that $\alpha x - y \geqslant 0$. Clearly if $x_i = 0$, for some i, then necessarily $y_i = 0$. □

Lemma 12. *Let G be a finite encoder. If G is not quasi-monomial, then there exists $1 \leqslant i \leqslant k$ such that, for all $1 \leqslant j \leqslant n$, we have*

$$\sum_{\substack{1 \leqslant l \leqslant k \\ l \neq i}} g_{l,j} = 0 \Rightarrow g_{i,j} = 0.$$

Proof of Lemma 12. Suppose that for all $i = 1,\dots,k$, there exists $1 \leqslant j \leqslant n$ such that $\sum_{l \neq i} g_{l,j} = 0$ and $g_{i,j} \neq 0$.

Let $1 \leqslant i \leqslant k$, and j satisfy the preceding condition. As $g_{l,j}$ is an element of $\hat{\mathbb{N}}$ for all $l = 1,\dots,n$, $\sum_{l \neq i} g_{l,j} = 0$ is equivalent to $g_{l,j} = 0$, for all $l \neq i$. As $g_{i,j} \neq 0$, this proves that G has an information position for i.

As this is true for all i, G is quasi-monomial. □

Proof of Theorem 10

- Let us suppose that G is quasi-monomial and weakly non-catastrophic. There exists a diagonal matrix $A = \mathrm{diag}(a_1,\dots,a_k)$, with $a_i \in \mathbb{N}$ for $1 \leqslant i \leqslant k$, a $k \times (n-k)$ matrix B, and a permutation matrix P of rank n such that $G = (A \mid B)P$. We can suppose, without loss of generality, that $P = I_n$.

As G is weakly non-catastrophic, according to Theorem 6, if $x=uG$ is an element of $\hat{\mathbb{N}}^n$, then u is an element of $\hat{\mathbb{Z}}^k$. But we have $x_i=a_iu_i$, for $i=1,\dots,k$ with $x_i\in\hat{\mathbb{N}}$ and $a_i\in\mathbb{N}$. So necessarily u_i is an element of $\hat{\mathbb{N}}$, for all i.

- Let us suppose that G is not quasi-monomial. Then, according to Lemma 12, there exists an index $1\leqslant i\leqslant k$ such that for all $j=1,\dots,n$

$$\sum_{\substack{1\leqslant l\leqslant k\\ l\neq i}} g_{l,j}=0\Rightarrow g_{i,j}=0.$$

So, using Lemma 11, there exists an integer α such that

$$\alpha\sum_{\substack{1\leqslant l\leqslant k\\ l\neq i}} g^{(l)}-g^{(i)}\geqslant 0,$$

where $g^{(l)}$ is the lth row of G.

Let $u=(u_l)_{1\leqslant l\leqslant k}$ be defined by

$$u_l=\begin{cases}-1 & \text{if } l=i,\\ \alpha & \text{otherwise.}\end{cases}$$

Then $x=uG$ is an element of $\hat{\mathbb{N}}^n$ and u is not an element of $\hat{\mathbb{N}}^k$.

So G is not strongly non-catastrophic. □

Definition 13. A strongly non-catastrophic encoder G is said to be *basic* if, when $x=uG$ is an element of $\mathbb{N}^n$, then necessarily u is an element of $\mathbb{N}^k$.

Definition 14. Given a $k\times n$ matrix M over K, we denote the gcd of its ith-order minors by Δ_i, for $i=1,\dots,k$, and, by convention, $\Delta_0=1$. Then, for $i=1,\dots,k$, the ith invariant factor of G is the element of K defined by $\gamma_i=\Delta_i/\Delta_{i-1}$.

Theorem 15. *Let G be a basic encoder. The following properties are satisfied*:

(1) *The invariant factors of G are all* 1, *i.e.* $\gamma_i=1$ *for all i.*

(2) *The gcd of the kth-order minors of G is* 1, *i.e.* $\Delta_k=1$.

(3) *G has a right inverse with coefficients in $\mathbb{Z}$.*

(4) *G is a sub-matrix of an $n\times n$ matrix invertible in $\mathbb{Z}$.*

Proof

Basic $\Rightarrow$ (2): As G is strongly non-catastrophic the gcd of its kth-order minors is a power of 2 (Theorem 6), say 2^l for some $l\geqslant 0$. Suppose that $l>0$. Let $\tilde{G}$ be the reduction of G modulo 2. As $l>0$ the rank of $\tilde{G}$ is strictly less than k. Then there exists a non-null vector $\tilde{U}$ such that $\tilde{U}\tilde{G}$ is null in GF(2). Let U be the canonical injection of $\tilde{U}$ in $\mathbb{N}$. Then every coefficient of the vector UG is a multiple of 2. Hence the vector $u=\frac{1}{2}U$ has at least one coefficient which is not in $\mathbb{N}$ and yet the vector $x=uG$ is in $\mathbb{N}^n$. So G is not basic.

(2) ⇔ (1): The product of the invariant factors of G is given by

$$\begin{aligned}\gamma_1\gamma_2\cdots\gamma_k &= \frac{\Delta_1}{\Delta_0}\frac{\Delta_2}{\Delta_1}\cdots\frac{\Delta_k}{\Delta_{k-1}}\\ &= \frac{\Delta_k}{\Delta_0}\\ &= \Delta_k.\end{aligned}$$

As every coefficient of G is in $\mathbb{N}$, the Δ_i are non-negative integers, and as the rank of G is at least 1, they are, in fact, positive. Moreover, by construction, Δ_{i-1} divides Δ_i for all $i \geqslant 1$, so the γ_i are also non-negative integers. Hence the γ_i are all 1 if and only if Δ_k is equal to 1.

(2) ⇒ (3): Denote by G_v the $k \times k$ sub-matrices of G and by Γ_v the corresponding minors, for $v = 1,\ldots,\binom{n}{k}$. Let, for $v = 1,\ldots,\binom{n}{k}$, K_v be the adjoint matrix of G_v, such that $G_vK_v = \Gamma_vI_k$. Inserting appropriately $n-k$ null rows in K_v, we obtain a matrix K'_v such that $GK'_v = \Gamma_vI_k$, for all v.

As the gcd of the Γ_v is 1, there exists some integers h_v such that $\sum_v h_v\Gamma_v = 1$. Let

$$K = \sum_v h_vK'_v.$$

Then K has its coefficients in $\mathbb{Z}$ and

$$\begin{aligned}GK &= \sum_v h_vGK'_v\\ &= \sum_v h_v\Gamma_vI_k\\ &= \left(\sum_v h_v\Gamma_v\right)I_k\\ &= I_k.\end{aligned}$$

(1) ⇔ (4): Denote by Φ the $k \times n$ matrix of the invariant factors of G (i.e. $\Phi = \operatorname{diag}(\gamma_1,\ldots,\gamma_k)$). Then there exists a $k \times k$ invertible matrix X and a $n \times n$ invertible matrix Y, having their coefficients in $\mathbb{Z}$, such that

$$XGY = \Phi.$$

As every invariant factor is 1, we have

$$G = A(I_k \quad 0_{k,n-k})B,$$

where $A = X^{-1}$ and $B = Y^{-1}$. If

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

where B_1 is a $k \times n$ matrix and B_2 is a $(n-k) \times n$ matrix, then $G = AB_1$. But, since A is invertible, the matrix

$$\begin{pmatrix} AB_1 \\ B_2 \end{pmatrix}$$

is also invertible and with coefficients in $\mathbb{Z}$.

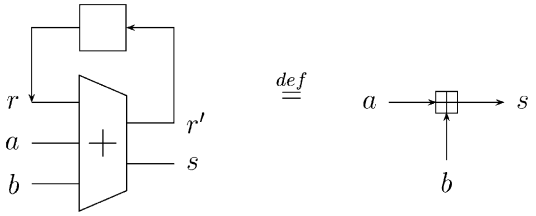

Fig. 1. Binary adder with carry.

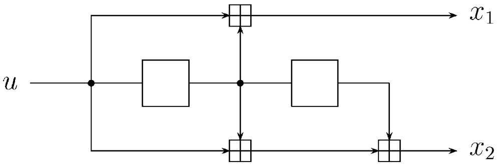

Fig. 2. Encoding circuit for the 2-adic encoder $G = (3, 7)$.

Conversely, if

$$B = \begin{pmatrix} G \\ H \end{pmatrix}$$

is invertible in $\mathbb{Z}$, then the equation $G = I_k \,(I_k \quad 0_{n-k})\, B$ shows that the invariant factors are all equal to 1. □

3.2. Physical realization

As for convolutional codes, it is possible to draw a blueprint for an actual physical device directly from a finite encoder of a given 2-adic code. In fact, as far as encoding circuits are concerned, the only difference between convolutional codes and 2-adic codes is the structure of binary adders. Adders for convolutional codes are simply XOR gates while adders for 2-adic codes are constituted by a three-entry modulo 2 adder with carry and a memory register which re-inject the carry at the next time index as shown in Fig. 1. In this figure $s = (a + b + r) \bmod 2$ and $r' = \lfloor (a + b + r)/2 \rfloor$. The encoding circuit for the 2-adic code generated by $G = (3, 7)$ is given in the example shown in Fig. 2.

Thus, to a given encoder correspond a unique encoding circuit and vice versa. The generic circuit of $(2, 1)$ codes is given in Fig. 3 (where a symbol inside a circle denotes a multiplicative gate). The parameters of this circuit are the following:

- m is a non-negative integer (constraint length of the code),
- for $i = 1, \ldots, m$, a_i is an element of GF(2) (feedback branch of the code),
- d^1, e^1 and, for $i = 1, \ldots, m$, c_i^1 are elements of GF(2) (forward branch for the first symbol) and

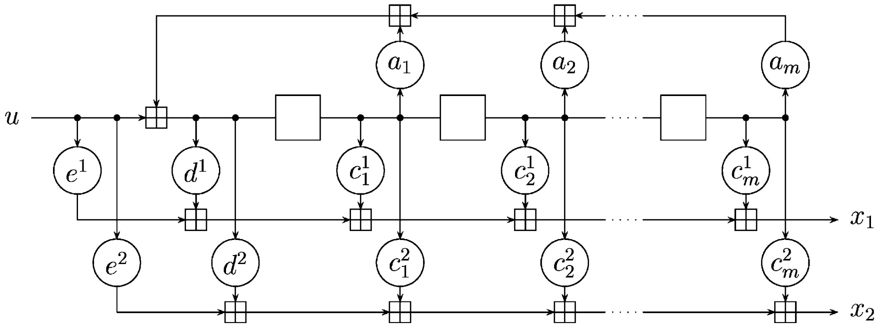

Fig. 3. Encoding circuit of a (2,1) code.

- d^2, e^2 and, for $i = 1,\ldots,m$, c_i^2 are elements of GF(2) (forward branch for the second symbol).

  Let

$$t = -1 + \sum_{i=1}^{m} a_i 2^i,$$

$$r_s = \sum_{i=1}^{m} c_i^s 2^i \quad \text{with } s = 1 \text{ or } 2.$$

Then the circuit presented in Fig. 3 is the encoding circuit of the encoder

$$G = \left(\frac{e^1 t - d^1 - r_1}{t} \quad \frac{e^2 t - d^2 - r_2}{t} \right).$$

(Note that G is a finite encoder if and only if $a_1 = \cdots = a_m = 0$, that is if there is no feedback branch.)

This construction can be easily extended to $(n,1)$ codes by adding more forward branches. Moreover, a (n,k) code, is, in fact, the superposition of k 1-dimensional circuits whose output vectors are added componentwise (with carry).

3.3. Transducers and trellis structure

Definition 16. Let V be any finite set, A and B two monoids (whose neutral elements are both denoted by 0). Let E be a subset of $V \times V \times A \times B$ and r an element of V. The triple $T = (V,E,r)$ is an (A,B)- *transducer* if

(1) For all v in V,

$$\biguplus_{(v,v',a,b)\in E} \{a\} = A.$$

(2) $(r,r,0,0)$ is an element of E.

Note that a transducer has a structure of labeled directed graph. An example of a $(\mathrm{GF}(2),\mathrm{GF}(2)^2)$-transducer is given in Fig. 4.

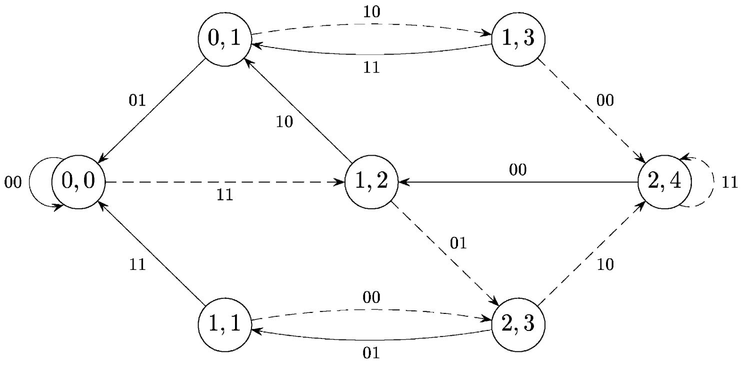

Fig. 4. Refined transducer for the 2-adic encoder $G=(3,5)$.

Definition 17. Let $T=(V,E,r)$ be an (A,B)-transducer. The *output function* associated with T is the function f_T defined by

$$f_T : A^{\mathbb{N}} \to B^{\mathbb{N}},$$

$$(a_0,a_1,\ldots) \mapsto (b_0,b_1,\ldots)$$

such that there exists v_1 in V such that $(r,v_1,a_0,b_0)\in E$ and, for all $i\geqslant 1$, there exists $v_{i+1}\in V$ such that $(v_i,v_{i+1},a_i,b_i)\in E$.

In Definition 16, property (1) insures that f_T is well defined over $A^{\mathbb{N}}$ and property (2) insures that if $f_T(a_0,a_1,\ldots)=(b_0,b_1,\ldots)$, then $f_T(0,a_0,a_1,\ldots)=(0,b_0,b_1,\ldots)$. Moreover, there exists a natural set isomorphism between any complete discrete valuation ring with residue field A and the set $A^{\mathbb{N}}$. Hence we will use the notation $f_T(\sum_{i\geqslant 0} a_i\pi^i)$ to denote $f_T(a_0,a_1,\ldots)$.

Theorem 18. *Let G be a finite encoder of an (n,k) 2-adic code $\mathscr{C}$. Then there exists a $(\mathrm{GF}(2)^k,\mathrm{GF}(2)^n)$-transducer T (associated to G) such that, for all u in $\mathbb{Q}_2^k$,*

$$uG = f_T(u).$$

Proof. We have seen in the preceding section that we can associate an encoding circuit to every finite encoder. This circuit is made of memory register and adders with carry. So the current state of the whole circuit is completely determined if we know the respective states of all the memory registers (including those inside the adders). Moreover for every state of the circuit and every input vector, we can compute the state in which the circuit will be during the next time index. In this way, we construct a transducer whose vertices are the possible states of the encoding circuit and in which two vertices are connected by an edge if and only if there exists an input vector transforming a vertex in the other.

As this transducer is just another formalization of the encoding circuit, it verifies the required property. □

In order to simplify the notations in the rest of this section, given a non-negative integer x, we will denote the quotient (resp. remainder) of x by 2 by $x^{\leftarrow}$ (resp. x°). (And we will apply them componentwise on vectors.)

Example 19. Let $G=(3,7)$. The encoding circuit associated with G is given in Fig. 2. There are 2 memory registers and 3 adders with carry, so a state of the circuit is given by five binary digits (r_1,r_2,c_1,c_2,c_3). If $u\in\mathrm{GF}(2)$ is encoded then the next state will be (u,r_1,c_1',c_2',c_3') with $c_1'=(u+r_1+c_1)^{\leftarrow}$, $c_2'=(u+r_1+c_2)^{\leftarrow}$, $c_3'=((u+r_1+c_2)^{\circ}+r_2+c_3)^{\leftarrow}$. The corresponding output is the vector $((u+r_1+c_1)^{\circ},(u+r_1+c_2+r_2+c_3)^{\circ})$. In this way, we construct a $(\mathrm{GF}(2),\mathrm{GF}(2)^2)$-transducer with 32 vertices and 64 edges whose initial state is $(0,0,0,0,0)$.

We can already see that this construction produces transducers which are far from being optimal. In the preceding example, we note that c_1 and c_2 must be equal. In fact, some states are never reached from the initial state (among others, the states with $c_1\neq c_2$). Some other states can be merged together because they have the same predecessors, successors and input/output values.

Theorem 20. *Let g be a positive integer. There exists a* $(\mathrm{GF}(2),\mathrm{GF}(2))$*-transducer T with g states such that $ug=f_T(u)$, for all $u\in\mathbb{Z}_2$.*

Proof. Let g be a positive integer and $u=\sum_{i\geqslant 0}u_i2^i$ an element of $\mathbb{Z}_2$. Let $(s_i)_{i\in\mathbb{N}}$ and $(x_i)_{i\in\mathbb{N}}$ be two sequences defined by

$$\left.\begin{array}{l} s_{i+1}=(s_i+u_ig)^{\leftarrow} \\ x_i=(s_i+u_ig)^{\circ} \end{array}\right. \quad \text{for } i\geqslant 0,$$

with $s_0=0$. Then

$$x_0=(u_0g)^{\circ}=(ug)^{\circ},$$

$$x_1=((u_0g)^{\leftarrow}+u_1g)^{\circ}=((ug)^{\leftarrow})^{\circ},$$

$$x_2=(((u_0g)^{\leftarrow}+u_1g)^{\leftarrow}+u_2g)^{\circ}=(((ug)^{\leftarrow})^{\leftarrow})^{\circ},$$

$$\vdots$$

As, for all $y\in\mathbb{N}$, $y=y^{\circ}+2(y^{\leftarrow})^{\circ}+4((y^{\leftarrow})^{\leftarrow})^{\circ}+\cdots$, then

$$\sum_{i\geqslant 0}x_i2^i=ug.$$

So, let $V=\{0,1,\ldots,g-1\}$ and $E=\{(s,s^{\leftarrow},0,s^{\circ})\mid s\in V\}\cup\{(s,(s+g)^{\leftarrow},1,(s+g)^{\circ})\mid s\in V\}$. Then $T=(V,E,0)$ is a $(\mathrm{GF}(2),\mathrm{GF}(2))$-transducer such that, for all $u\in\mathbb{Z}_2$,

$$ug=f_T(u). \qquad \square$$

Using the preceding theorem, we can construct a transducer for every vector $G=(g_1,\ldots,g_n)$ of positive integers. Let $V=\prod_{i=1}^{n}\{0,1,\ldots,g_i-1\}$ and $E=\{(s,(s+$

$uG)^{\leftarrow}, u, (s + uG)^{\circ}) \mid s \in V,\ u \in \{0,1\}\}$. The transducer $(V, E, 0)$ is called the *product transducer* of G. Once again this transducer is too big, in the sense that some vertices cannot be reached from the initial state. So let $V^{\star}$ be the subset of V containing only vertices which are successors of the 0 state and $E^{\star} = \{(s, s', i, o) \in E \mid s \in V^{\star}\}$. The transducer $(V^{\star}, E^{\star}, 0)$ is called the *refined transducer* of G. Clearly the number of vertices of the product transducer is $\prod_{i=1}^{n} g_i$, but the numbers of vertices of the refined transducer is much less than this product. We have the following result.

Proposition 21. *Let $G = (g_1, \ldots, g_n)$ be a vector of n positive integers. Let $T = (V^{\star}, E^{\star}, 0)$ be the refined trellis of G. Then*

$$|V^{\star}| \leqslant \left(\sum_{i=1}^{n} g_i \right) - n + 1. \tag{1}$$

Proof. Let $(V, E, 0)$ be the product transducer of G. For an element $s = (s_1, \ldots, s_n)$ of V we say that s is admissible if

$$\text{for all } 1 \leqslant i < j \leqslant n, \quad -g_i < s_j g_i - s_i g_j < g_j. \tag{2}$$

Denote the set of admissible elements by $\mathscr{A}$. We will show that $V^{\star}$ is a subset of $\mathscr{A}$ and that the number of elements of $\mathscr{A}$ verifies (1).

- Let s be an admissible element of V and $1 \leqslant i < j \leqslant n$. Then

  $$-g_i < (2s_j^{\leftarrow} + s_j^{\circ})g_i - (2s_i^{\leftarrow} + s_i^{\circ})g_j < g_j.$$

  This is equivalent to

  $$-g_i - s_j^{\circ} g_i + s_i^{\circ} g_j < 2s_j^{\leftarrow} g_i - 2s_i^{\leftarrow} g_j < g_j + s_i^{\circ} g_j - s_j^{\circ} g_i.$$

  As $-2g_i \leqslant -g_i - s_j^{\circ} g_i + s_i^{\circ} g_j$ and $g_j + s_i^{\circ} g_j - s_j^{\circ} g_i \leqslant 2g_i$ (s_i° and s_j° are 0 or 1), then $-g_i < s_j^{\leftarrow} g_i - s_i^{\leftarrow} g_j < g_j$.Moreover, as s is admissible, we have

  $$-g_i < (s_j + g_j)g_i - (s_i + g_i)g_j < g_j.$$

  Using the same decomposition as above, we have

  $$-g_i < (s_j + g_j)^{\leftarrow} g_i - (s_i + g_i)^{\leftarrow} g_j < g_j.$$

  Hence if s is admissible then $s^{\leftarrow}$ and $(s + G)^{\leftarrow}$ are also admissible. So, as 0 clearly verifies (2), every element of $V^{\star}$ is admissible.

- Let s be an admissible element of V and $e = (e_1, \ldots, e_n)$ be a non-zero vector of $\mathbb{Z}^n$ such that $\sum_{i=1}^{n} e_i = 0$. Then there exists an index i such that $e_i > 0$ and an index j such that $e_j < 0$. Let's suppose that $-g_i < (s_j + e_j)g_i - (s_i + e_i)g_j < g_j$. This is equivalent to

  $$-g_i - e_j g_i + e_i g_j < s_j g_i - s_i g_j < g_j + e_i g_j - e_j g_i.$$

  As s verifies (2), we have

  $$-g_i \leqslant -g_i - e_j g_i + e_i g_j,$$
  $$g_j + e_i g_j - e_j g_i \leqslant g_j.$$

Hence $e_i g_j - e_j g_i = 0$ which, as $e_i > 0$ and $e_j < 0$, is impossible. So $s + e$ is not admissible. In other words, given an positive integer t there exists at most one vector of $V^\star$ whose coordinates sum to t. So,

$$|V^\star| \leqslant 1 + \sum_{i=1}^{n} (g_i - 1). \qquad \square$$

Note that bound (1) can still be refined. For example, in the case $n = 2$, let $G = (g_1, g_2)$, with g_1 and g_2 being two positive integers. Denote the gcd of g_1 and g_2 by d. We can see that, for $u = 1, \ldots, d$, there is not any element (s_1, s_2) in $\mathscr{A}$ with $s_1 + s_2 = u(g_1 + g_2)/d - 1$. So, in this case $|V^\star| \leqslant g_1 + g_2 - d$. This can be generalized in the case $n > 2$ as

$$|V^\star| \leqslant \sum_{t=1}^{n} (-1)^{t+1} \left(\sum_{i_1=1}^{n-t+1} \sum_{i_2=i_1+1}^{n-t+2} \cdots \sum_{i_t=i_{t-1}+1}^{n} \gcd(g_{i_1}, \ldots, g_{i_t}) \right).$$

(With the natural convention $\gcd(x) = x$ for every x.)

Moreover, we conjecture that this bound is exact, that is:

Conjecture 22. *Let $G = (g_1, \ldots, g_n)$ be a vector of n positive integers. Let $T = (V^\star, E^\star, r)$ be the refined trellis of G. Then*

$$|V^\star| = \sum_{t=1}^{n} (-1)^{t+1} \left(\sum_{i_1=1}^{n-t+1} \sum_{i_2=i_1+1}^{n-t+2} \cdots \sum_{i_t=i_{t-1}+1}^{n} \gcd(g_{i_1}, \ldots, g_{i_t}) \right).$$

Moreover, if the g_i's are pairwise coprime, then

$$|V^\star| = \sum_{i=1}^{n} g_i - n + 1.$$

Note 1. Since the original redaction of this paper, I proved this conjecture. In fact, a transducer is, among other things, a Markov chain having an unique ergodic class, so we can study its stationary probability distribution vector. Using this vector, we can show that an element of V belongs to $V^\star$ if and only if it is admissible. The result then follows.

Example 23. Fig. 4 shows as an example the refined transducer for the 2-adic code generated by $G = (3, 5)$. In this figure an arrow denotes a state transition, it is plain if the input is 0 and dashed if the input is 1 and its label is the corresponding output vector. The set of vertices is a subset of $\{0, \ldots, 2\} \times \{0, \ldots, 4\}$ with 7 elements.

For a given code, the existence of a transducer implies the existence of a trellis describing this code for any length. So, in order to decode this code, we can use any trellis-oriented algorithm, for example the Viterbi algorithm.

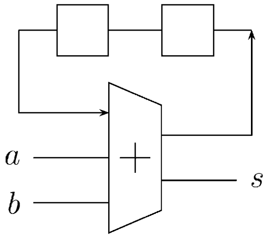

Fig. 5. Adder with carry in $\mathbb{Q}_2[X]/(X^2-2)$.

4. Codes over extensions of the 2-adic field

Consider the polynomial $P(X)=X^e-2$ in $\mathbb{Q}_2[X]$, for $e\geqslant 2$. Let θ be a root of $P(X)$ and $L=\mathbb{Q}_2(\theta)$. As $2=\theta^e$, 2 is nomore prime in L, and so is not an uniformizer. However, θ is such an uniformizer. Hence $L=\{\sum_{i\geqslant r}x_i\theta^i \mid r\in\mathbb{Z}, x_i\in \mathrm{GF}(2)\}$. We say that $L/\mathbb{Q}_2$ is a totally ramified extension of degree e. (More generally, this is true as soon as $P(X)=X^e+a_{e-1}X^{e-1}+\cdots+a_1X+a_0$ is an Eisenstein polynomial, i.e. $2\mid a_i$, for $0\leqslant i<e$ and $4\nmid a_0$.)

Note that we can also write $L=\mathbb{Q}_2+\theta\mathbb{Q}_2+\cdots+\theta^{e-1}\mathbb{Q}_2$ but, in this case, the sequence is e-interleaved. So, if $P(X)=X^e-2$ then the carry computed at time t is in fact added at time $t+e$. This means that an adder for L contains e memory registers (see Fig. 5 in the case $e=2$).

Moreover, the techniques used for computing the transducer in Section 3.3 can still be used for $e>1$, with, for x in L, $x^{\leftarrow}$ (resp. x°) denoting the quotient (resp. remainder) of x by θ.

5. Some remarks about 2-adic codes

We give here some remarks about 2-adic codes concerning their difference against convolutional codes.

- *Non* GF(2)-*linearity*. Due to the change of characteristic between GF(2) and $\mathbb{Q}_2$, a 2-adic code is linear over $\mathbb{Q}_2$ but not over GF(2). So we cannot represent a 2-adic encoder as a semi-infinite matrix with coefficients over GF(2) (as it can be done for convolutional codes).
- *Distance properties.* The lack of GF(2)-linearity of 2-adic codes implies, among other things, that the Hamming distance is not directly induced by the Hamming weight (i.e. $d_H(x,y)\neq w_H(x-y)$, for some x and y). Clearly, a notion of weight enumerator can be defined (directly or on the trellis structure) but its usefulness is not evident. So we must give a notion of distance enumerator by iterating over every pair of codewords. This is mathematically possible but obviously impossible to compute in practice. Worse, although a crude approximation may be possible to compute, it is likely unuseful.

- *Performances*. For the moment, we made a very small number of simulations, in fact, just enough to validate the encoding and decoding algorithms. It seems that the performance of a 2-adic code is very close to that of a convolutional code with the same parameters. Note that we just compared convolutional codes which are known to be optimal with 2-adic codes chosen more or less at random (but with comparable decoding complexity). Nevertheless, for a 2-adic code, the sequence of carries depends on the message (more precisely, on the distribution of the 1's in the message) and is transmitted as part of the codeword. This suggest a better decoding power.

6. Conclusion

We have seen in this paper that we can define codes over p-adic fields and that they can be used in the same way as convolutional codes. They can be encoded and decoded easily, using already known techniques. Moreover, these codes can, a priori, be used everywhere the convolutional codes are used (alone or, for example, as constituents of turbo-codes).

Nevertheless, a great amount of work is still to be done, notably to give a mathematical notion of what is a "good" 2-adic code. It also seems that codes defined over a ramified extension of a p-adic field may be of particular interest (the carry sequence is then more or less interleaved between the symbols).

Uncited References

[2–5,7,11]

Acknowledgements

I thank T. Berger and P. Piret for helpful conversations we had on this subject and the referees for their comments on the original version of this paper.

References

[1] A.R. Calderbank, N.J.A. Sloane, Modular and p-adic cyclic codes, Des. Codes Cryptogr. 6 (1995) 21–35.

[2] H. Cohen, A Course in Computational Algebraic Number Theory, Graduate texts in Mathematics, Springer, Berlin, 1995.

[3] F.Q. Gouvea, p-adic Numbers, Springer, Berlin, 1991.

[4] R. Hammons, V. Kumar, A.R. Calderbank, N.J.A. Sloane, P. Solé, The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes, IEEE Trans. Inform. Theory, 1994.

[5] H. Hasse, Number Theory, Springer, Berlin, 1980.

[6] R.J. McEliece, The algebraic theory of convolutional codes, in: Handbook of Coding Theory, 1996.

[7] P. Piret, Convolutional Codes: an Algebraic Approach, MIT Press, Cambridge, MA, 1988.
[8] P. Ribenboim, L'Arithmétique des corps, Hermann, Paris, 1972.
[9] J.P. Serre, Local Fields, Graduate Texts in Mathematics, Vol. 67, Springer, Berlin, 1979.
[10] P. Solé, Open problem 2: cyclic codes over rings and p-adic fields, in: G. Cohen, J. Wolfmann (Eds.), Coding Theory and Applications, Lecture Notes in Computer Science, Vol. 388, Springer, New York, 1988, p. 329.
[11] J. Vuillemin, On circuits and numbers, Rapport, Digital Equipment Corp. (PRL), 1993.