# Introducción a la Teoría de Códigos

Guillermo Morales-Luna
Centro de Investigación y Estudios Avanzados del IPN
(CINVESTAV-IPN)
gmorales@cs.cinvestav.mx

9 de mayo de 2010

**Resumen**

Presentamos las nociones fundamentales de Teoría de Códigos. Hemos querido ponderar un enfoque algebraico, pero también computacional, procurando describir los procedimientos de manera que cualquier lector pueda programarlos en el lenguaje de su preferencia. Tras una breve introducción, damos las definiciones básicas de códigos, y de procedimientos de decodificación de tipo "instantáneo", luego revisamos los códigos binarios, y pasamos después a los lineales donde enfatizamos el significado geométrico de las nociones involucradas; vemos también algunos procedimientos para obtener nuevos códigos partiendo de otros ya construídos previamente. Como códigos particulares, presentamos los códigos de Reed-Muller y los códigos cíclicos, entre los que se cuentan los de Golay. Presentamos también los códigos de Reed-Solomon (siguiendo dos enfoques, el primero de tipo "práctico" y el segundo de tipo algebraico) y los códigos BCH.

Estas notas se encuentran en un constante proceso de (re-)elaboración.

## Contenido

# 1 Introducción

En esta breve introducción a la Teoría de Códigos seguiremos cercanamente la exposición de Adamek [1] así como el libro de van Lint [13]. Las nociones básicas de Teoría de Números, desarrolladas teniendo de fondo a sus aplicaciones en Comunicaciones, están claramente expuestas en el libro de Schroeder [11]. Niederreiter es un autor muy conocido, ya clásico, en Teoría de Campos, y su texto en códigos [9] es sumamente recomendable, y lo mismo puede decirse de Washington, cuyo libro fue muy conocido desde su primera edicón [15] y la segunda fue sustancialmente actualizada [12]. El manual compilado por Huffman y Brualdi [8] ya es indispensable para los especialistas de esta área. Textos más recientes son [5, 6, 7, 10].

# 2 Codificación

## 2.1 Funciones de codificación

Sea $A$ un alfabeto finito y sea $C$ un alfabeto llamado *de código*. $C^+$ denota al conjunto de palabras no-vacías con símbolos en $C$. Una *codificación* es una función $\gamma : A \to C^+$. La imagen de $\gamma$ se dice ser el conjunto de *palabras de código* (*codewords*). $\gamma$ se extiende, mediante la mera concatenación de valores, a una función $\gamma^* : A^* \to C^+$, donde $A^* = \bigcup_{n\geq 0} A^n = A^+ \cup \{\text{nil}\}$, haciendo:

$$\gamma^*(\text{nil}) = \text{nil} \qquad , \qquad \forall \sigma \in A^*, a \in A : \ \gamma^*(\sigma a) = \gamma^*(\sigma)\,\gamma(a).$$

La codificación se dice ser de una *decodificación única* si $\gamma^*$ es una función inyectiva.

**Ejemplo 2.1 (4-en-8)** *Codifiquemos un alfabeto de 70 símbolos mediante un alfabeto consistente de tiras de 8 bits, 4 de los cuales están prendidos y los otros 4 apagados.*

Sea $A$ un alfabeto de 70 símbolos, por ejemplo, $A$ podría ser el alfabeto que consta de las 26 letras minúsculas, las 26 mayúsculas, los 10 dígitos decimales y 8 símbolos de puntuación (entre los cuales está el "blanco"). Sea $C = \{0, 1\}$ el alfabeto de sólo dos símbolos y sea $C_e^8$ el conjunto de palabras de 8 bits en las que exactamente hay 4 bits con valor 1. Es claro que $\text{card}\,(C_e^8) = \binom{8}{4} = 70$ y que el conjunto $C_e^8$ está ordenado con el orden usual de los enteros cuando cada palabra se lee como la representación en base-2 de un entero: El primer elemento es pues $'00001111'$ y el último $'11110000'$. Sea $\gamma : A \to C_e^8$ la función que asocia el $i$-ésimo elemento de $A$ con el $i$-ésimo elemento de $C_e^8$. Una palabra de longitud $k$ en $A$ quedará entonces codificada por $8k$ bits. Cada bloque de 8 bits contiguos, de la forma $\varepsilon_{8\ell-7}\varepsilon_{8\ell-6}\cdots\varepsilon_{8\ell-1}\varepsilon_{8\ell}$, con $1 \leq \ell \leq k$, contiene exactamente 4 1's. Naturalmente, de las $2^8 = 256$ palabras de 8 bits, hay $2^8 - \binom{8}{4} = 186$ que no son palabras de código. La decodificación es natural. Cada uno de los bloques de 8 bits contiguos determina un símbolo de $A$. □

**Observación 2.1 (Terminología)** *En la literatura técnica es común nombrar a una función de codificación como un* código*, e incluso, a su imagen, o sea al conjunto de palabras de código, se le llama también* código*. Así pues la palabra* código *tiene prácticamente cuatro connotaciones: el esquema $(\gamma, A, C)$, la función $\gamma$, el valor $\gamma(a)$ asociado a cada símbolo $a \in A$, y la propia imagen $\gamma(A)$. De estas cuatro, la mayormente supuesta en la* Teoría de Códigos *es la cuarta.*

Si existe una $k \in \mathbb{N}$, $k > 0$, tal que $\gamma(A) \subset C^k$, entonces el código se dice ser *de bloques*: cada símbolo se codifica mediante un bloque de $k$ símbolos.

**Ejemplo 2.2 (ASCII)** *El código ASCII (*American Standard Code for Information Interchange*) es un código por bloques de 8 bits (con una connotación inicial de 7 bits más 1 de paridad).*

Por otro lado, el código se dice ser *instantáneo* si no ocurre que el código de un símbolo sea un prefijo del código de otro símbolo:

$$\forall a_1, a_2 \in A : \ [[\exists \tau \in C^* : \gamma(a_2) = \gamma(a_1)\tau] \ \Rightarrow \ a_1 = a_2]\,.$$

| | | | |
|---|---|---|---|
| A | pre | N | rpe |
| B | rpppe | O | rrre |
| C | rprpe | P | prrpe |
| D | rppe | Q | rrpre |
| E | pe | R | prpe |
| F | pprpe | S | pppe |
| G | rrpe | T | re |
| H | ppppe | U | ppre |
| I | ppe | V | pppre |
| J | prrre | W | prre |
| K | rpre | X | rppre |
| L | prppe | Y | rprre |
| M | rre | Z | rrppe |

(a)

| | | | |
|---|---|---|---|
| E | pe | O | rrre |
| T | re | H | ppppe |
| I | ppe | V | pppre |
| A | pre | F | pprpe |
| N | rpe | L | prppe |
| M | rre | P | prrpe |
| S | pppe | J | prrre |
| U | ppre | B | rpppe |
| R | prpe | X | rppre |
| W | prre | C | rprpe |
| D | rppe | Y | rprre |
| K | rpre | Z | rrppe |
| G | rrpe | Q | rrpre |

(b)

Recuadro 1: Código Morse. (a) En orden alfabético. (b) En orden de frecuencias en el idioma inglés.

**Ejemplo 2.3 (Código Morse)** *Sea Latino el alfabeto de las 26 letras mayúsculas y sea* $Cd = \{\texttt{p},\texttt{r},\texttt{e}\}$ *el afabeto consistente de tres símbolos:* punto, raya, espacio. *El código es instantáneo y se muestra en la tabla 1*

El Código Morse es de decodificación única: ya que la codificación de cada símbolo termina con "e" (espacio), entre dos espacios contiguos cualesquiera se decodifica un único símbolo. También, al asignar códigos más cortos a símbolos más frecuentes, podrá esperarse que en el Código Morse los códigos de palabras serán de longitudes "cortas". □

Los códigos instantáneos deben su nombre a un procedimiento "natural" para decodificarlos: Leyéndolos de izquierda a derecha, cada vez que se reconoce el código de un símbolo se escribe ese símbolo y se reinicia el proceso de reconocimiento.

## 2.2 Decodificación única

**Teorema 2.1 (Kraft)** *Sea* $A = \{a_i\}_{i=0}^{n-1}$ *un alfabeto de* $n$ *símbolos y sea* $L = \{\ell_i\}_{i=0}^{n-1}$ *un conjunto de* $n$ *números naturales distintos de cero. Existe un código instantáneo, utilizando un alfabeto de código* $C$ *de* $k$ *símbolos, de manera que cada* $a_i$ *corresponda a una palabra de longitud* $\ell_i$ *cuando y sólo cuando se cumpla la* desigualdad de Kraft

$$\sum_{i=0}^{n-1} k^{-\ell_i} \leq 1. \tag{1}$$

En efecto, haciendo un renombramiento de símbolos si fuera necesario, podemos suponer que $L$ es no-decreciente: $1 \leq \ell_0 \leq \ell_1 \leq \cdots \leq \ell_{n-1}$.

Supongamos que existe un código instantáneo $\gamma : A \to C^*$, $a_i \mapsto \gamma(a_i)$, tal que

$$\operatorname{long}(\gamma(a_i)) = \ell_i \ , \ 0 \leq i \leq n-1.$$

Al ser el código instantáneo, para cada $i > 0$, ningún $\gamma(a_j)$, con $j < i$, puede ser un prefijo de $\gamma(a_i)$. Así pues, se tiene exactamente $k^{\ell_i} - \sum_{j=0}^{i-1} k^{\ell_i - \ell_j}$ posibilidades de elegir $\gamma(a_i)$. En consecuencia,

$$\forall i > 0 : \ k^{\ell_i} \geq \sum_{j=0}^{i-1} k^{\ell_i - \ell_j} + 1. \tag{2}$$

En particular, para $i = n-1$, al dividir ambos miembros de la desigualdad (2) entre $k^{\ell_{n-1}}$ resulta la desigualdad de Kraft (1).

Recíprocamente suponiendo que vale (1) entonces puede verse consecutivamente que valen las relaciones (2). El código instantáneo se construye de manera sucesiva: Para $\gamma(a_0)$ elíjase una de las $k^{\ell_0}$ palabras posibles de longitud $\ell_0$. Para cada $i > 0$ habiendo construído $\gamma(a_j)$, con $j < i$, elíjase como $\gamma(a_i)$ a una palabra de longitud $\ell_i$ que no tenga como prefijo a ninguna $\gamma(a_j)$. Por la desigualdad (2), esto último siempre es posible. □

**Teorema 2.2 (McMillan)** *Cualquier código de decodificación única satisface la desigualdad de Kraft.*

Sea $A = \{a_i\}_{i=0}^{n-1}$ un alfabeto de $n$ símbolos y sea $\gamma : A \to C^*$ una codificación de decodificación única, sobre un alfabeto $C$ de $k$ símbolos. Sea $\ell_i = \operatorname{long}(\gamma(a_i))$, $i = 0, \ldots, n-1$. Escribamos $\lambda = \sum_{i=0}^{n-1} k^{-\ell_i}$ y mostremos que, en efecto, $\lambda \leq 1$. Para esto deberemos demostrar que la sucesión $\left(\frac{\lambda^\nu}{\nu}\right)_{\nu \geq 1}$ está acotada superiormente (obviamente vale la implicación: $\lambda > 1 \Rightarrow \lim_{\nu \to +\infty} \frac{\lambda^\nu}{\nu} = +\infty$). Calculemos entonces las potencias de $\lambda$:

$$\begin{aligned}
\lambda^2 &= \left(\sum_{i_0=0}^{n-1} k^{-\ell_{i_0}}\right)\left(\sum_{i_1=0}^{n-1} k^{-\ell_{i_1}}\right) = \sum_{i_0,i_1=0}^{n-1} k^{-(\ell_{i_0}+\ell_{i_1})} \\
\lambda^3 &= \left(\sum_{i_0,i_1=0}^{n-1} k^{-(\ell_{i_0}+\ell_{i_1})}\right)\left(\sum_{i_2=0}^{n-1} k^{-\ell_{i_2}}\right) = \sum_{i_0,i_1,i_2=0}^{n-1} k^{-(\ell_{i_0}+\ell_{i_1}+\ell_{i_2})} \\
&\vdots \quad \vdots \quad \vdots \\
\lambda^\nu &= \sum_{i_0,i_1,i_2,\ldots,i_{\nu-1}=0}^{\nu-1} k^{-(\ell_{i_0}+\ell_{i_1}+\ell_{i_2}+\cdots+\ell_{i_{\nu-1}})} \\
&\vdots \quad \vdots \quad \vdots
\end{aligned}$$

Ahora bien, para cada entero $\mu \in \mathbb{N}$ que se realice como la suma de $\nu$ logitudes, sea

$$I_{\mu\nu} = \left\{(i_0, i_1, i_2, \ldots, i_{\nu-1}) | \mu = \ell_{i_0} + \ell_{i_1} + \ell_{i_2} + \cdots + \ell_{i_{\nu-1}}\right\}$$

el conjunto de formas de expresar a $\mu$ como una tal suma y sea

$$S_{\mu\nu} = \{\sigma \in A^\nu | \gamma^*(\sigma) \in C^\mu\} = \{a_{i_0} a_{i_1} a_{i_2} \cdots a_{i_{\nu-1}} | (i_0, i_1, i_2, \ldots, i_{\nu-1}) \in I_{\mu\nu}\}$$

el conjunto de palabras de longitud $\nu$ en $A$ que quedan codificadas por palabras de longitud $\mu$ en $C$. Ya que el código es de decodificación única, se tiene

$$\operatorname{card}(I_{\mu\nu}) = \operatorname{card}(S_{\mu\nu}) \leq \operatorname{card}(C^\mu) = k^\mu.$$

Sea $\ell = \max[\ell_i]_{i=0}^{n-1}$ la mayor de las longitudes de códigos de símbolos. Naturalmente, el mayor de los valores $\mu$ tales que $I_{\mu\nu} \neq \emptyset$ es precisamente $\nu\ell$. En consecuencia:

$$\lambda^\nu = \sum_{\mu=1}^{\nu\ell} \sum_{(i_0,i_1,\ldots,i_{\nu-1}) \in I_{\mu\nu}} k^{-\mu} \leq \sum_{\mu=1}^{\nu\ell} k^\mu \cdot k^{-\mu} = \sum_{\mu=1}^{\nu\ell} 1 = \nu\ell,$$

de donde resulta $\frac{\lambda^\nu}{\nu} \leq \ell$, y esto ocurre para cualquier $\nu$. □

Como una consecuencia directa de ambos teoremas, resulta el

**Corolario 2.1** *Todo código de decodificación única da origen a códigos instantáneos que preservan las longitudes de los códigos asociados a los símbolos.*

## 2.3 Algoritmo de Sardinas-Patterson

Sea $\gamma : A \to C^+$ una función de codificación. Recordamos que confundiremos voluntariamente $\gamma^* : A^* \to C^*$ con $\gamma$. La función $\gamma$ es de decodificación única si rige la implicación siguiente:

$$\gamma(\alpha_0)\cdots\gamma(\alpha_{\mu-1}) = \gamma(\beta_0)\cdots\gamma(\beta_{\nu-1}) \implies \mu = \nu \ \& \ [\forall i < \mu : \ \alpha_i = \beta_i] \tag{3}$$

cualesquiera que sean los símbolos $\alpha_0, \ldots, \alpha_{\mu-1}\,;\, \beta_0, \ldots, \beta_{\nu-1} \in A$.

Sea $\Gamma = \gamma(A) \subset C^+$ la imagen bajo la función de codificación $\gamma$ del alfabeto $A$. Recursivamente, se define $\Gamma^0 = \{\text{nil}\}$ y $\Gamma^{k+1} = \Gamma^k\Gamma$. Para cada $k \in \mathbb{N}$, $\Gamma^k$ es pues la familia de palabras en $C^+$ consistente de códigos de palabras de longitud $k$ en $A$.

Si $\gamma$ es de decodificación única, entonces la relación (3) implica que toda palabra en $\Gamma^k$ se expresa de manera única como la concatenación de $k$ palabras en $\Gamma$. Sintetizaremos esta última propiedad, diciendo que $\Gamma^k$ es de *factorización única* sobre $\Gamma$.

**Proposición 2.1** *Si $\gamma$ es de decodificación única, entonces para cualesquiera $n, k \in \mathbb{Z}^+$, $\Gamma^{kn}$ es de factorización única sobre $\Gamma^k$.*

Se define también $\Gamma^* = \bigcup_{k\in\mathbb{N}} \Gamma^k$.

**Proposición 2.2 (Criterio de Schützenberger)** *Considérese las siguientes tres condiciones:*

1. *$\gamma$ es de decodificación única.*
2. $\Gamma \cap \left[\bigcup_{k\geq 2} \Gamma^k\right] = \emptyset$
3. $[(\sigma\Gamma^*) \cap \Gamma^* \neq \emptyset \ o \ (\Gamma^*\sigma) \cap \Gamma^* \neq \emptyset] \implies \sigma \in \Gamma$.

*Entonces vale la equivalencia:*

$$1. \iff 2. \ \& \ 3. \tag{4}$$

En efecto, veamos cada una de las implicaciones requeridas. Por un lado:
*1.* $\implies$ *2.* Es claro que si hubiera una palabra $\rho \in \Gamma \cap \left[\bigcup_{k\geq 2} \Gamma^k\right]$, ella tendría más de una factorización y en consecuencia $\gamma$ no podría ser de decodificación única.

*1.* $\implies$ *3.* Supongamos que hubiese una palabra $\tau \in (\sigma\Gamma^*) \cap \Gamma^*$. Existen entonces $k_0, k_1 \in \mathbb{N}$ tales que $\rho \in \sigma\Gamma^{k_0} \cap \Gamma^{k_1}$. Por la descomposición única *a fortiori* $k_1 = k_0 + 1$ y $\sigma \in \Gamma$.

Por otro lado:
*2. & 3.* $\implies$ *1.* Si acaso $\gamma(\alpha_0)\cdots\gamma(\alpha_{k_0-1}) = \gamma(\beta_0)\cdots\gamma(\beta_{k_1-1})$, entonces bien $\gamma(\alpha_0)$ es un prefijo de $\gamma(\beta_0)$ o al revés. Supongamos lo primero, entonces $\gamma(\beta_0) = \gamma(\alpha_0)\sigma$, para alguna palabra $\sigma$. Por la condición *3.* $\sigma$ debe estar en $\Gamma$. Por la condición *2.* se ha de tener $\gamma(\alpha_0) = \gamma(\beta_0)$ y $\alpha_0 = \beta_0$. Continuando de la misma forma para el resto de la palabra resulta la condición *1.* □

Sea $\gamma : A \to C^+$ una función de codificación y $\Gamma = \gamma(A) \subset C^+$ su imagen. Definimos, iterativamente

$$\begin{aligned} \Gamma_1 &= \{\sigma \in A^+ |\ \exists \tau \in \Gamma : \ \tau\sigma \in \Gamma\} \\ \Gamma_{k+1} &= \{\sigma \in A^+ |\ \exists j < k \ \left[\exists \tau \in \Gamma^j : \ \tau\sigma \in \Gamma^k\right] \ \text{o} \ \left[\exists \tau \in \Gamma^k : \ \tau\sigma \in \Gamma^j\right]\} \end{aligned}$$

El conjunto $\Gamma_1$ consiste pues de los *restos* que le hacen falta a palabras en el código $\Gamma$ para convertirse en otras palabras en el código. Similarmente, $\Gamma_k$ consiste de restos para que palabras previas se conviertan en códigos de palabras de longitud $k-1$, o bien de restos para que códigos de palabras de longitud $k-1$ se conviertan en palabras previas.

**Proposición 2.3** *Existen $i, j \in \mathbb{N}$, $i < j$, tales que $\Gamma_i = \Gamma_j$. Es decir, a la larga, la sucesión de conjuntos de restos $(\Gamma_k)_k$ ha de ser periódica.*

En efecto, veamos primero que las longitudes de los restos están acotadas. Sea $n$ la longitud mayor de las palabras de código: $n = \max\{\text{long}(\gamma(\alpha))|\ \alpha \in A\}$. Por inducción en $k$ se ve de manera directa:

$$\forall k \in \mathbb{Z}^+\ \forall \sigma \in \Gamma_k :\ \text{long}(\sigma) \leq n.$$

Así, cada $\Gamma_i$ sólo tiene un número finito de posibilidades de ser formado. De donde resulta la proposición. □

Naturalmente:

**Proposición 2.4** *El código $\gamma$ es de decodificación única si y sólo si*

$$\forall k \in \mathbb{Z}^+ :\ \Gamma \cap \Gamma_k = \emptyset. \tag{5}$$

Ahora bien, la condición $\Gamma \cap \Gamma_k \neq \emptyset$ es claramente equivalente a que nil $\in \Gamma_{k+1}$. Por tanto, si acaso nil $\in \Gamma_k$ para algún $k$ entonces $\gamma$ no puede ser de decodificación única.

Esto proporciona el *algoritmo de Sardinas-Patterson* para decidir si un código $\gamma$ es de decodificación única:

1. $\Gamma := \gamma(A) \subset C^+$ ;
2. $i := 1$ ; calcúlese $\Gamma_i$ ;
3. mientras que nil $\notin \Gamma_i$ hágase
   (a) $i++$ ;
   (b) calcúlese $\Gamma_i$ ;
   (c) si $(\exists j < i :\ \Gamma_j == \Gamma_i)$ entonces
   decídase '$\gamma$ es de decodificación única'
4. decídase '$\gamma$ no es de decodificación única'

(si ocurre que en alguna iteración nil $\in \Gamma_i$, entonces la palabra $\tau$ en $\Gamma_j \cap \Gamma_i$, con $j < i$, que mostrara tal hecho ha de poseer más de una factorización).

# 3 Códigos de Huffman

Los códigos de Huffman son instantáneos y asocian las cadenas más cortas a los caracteres más frecuentes. Sea $A$ un alfabeto en el que a cada símbolo $a \in A$ se le ha asociado un valor, digamos $f(a)$, llamado *peso* de $a$. Por ejemplo, para un "corpus" dado $\sigma \in A^*$, para cada $a \in A$ se cuenta el número de apariciones de $a$ en $\sigma$ para obtener el valor $c_\sigma(a)$ y se toma la "frecuencia" $f(a) = c_\sigma(a)/\text{long}(\sigma)$. Podemos pues suponer que para cada $a \in A$, $0 \leq f(a) \leq 1$ y $\sum_{a\in A} f(a) = 1$.

## 3.1 Códigos de Huffman binarios

Para construir los códigos binarios se procede de una manera arbórea:

1. Sea $\mathcal{A}$ la lista de elementos de la forma $(a, f(a))$, donde $f(a)$ es el peso del carácter $a$. Acaso mediante un renombramiento de los caracteres se puede suponer que los pesos están ordenados de manera no-decreciente. En este momento, $\mathcal{A}$ consta del follaje de un árbol a construirse.
2. Inicialmente, sea $\mathcal{T}$, el árbol a construirse, un árbol binario vacío.
3. En tanto sea posible repítase el ciclo siguiente:
   (a) Sáquese de $\mathcal{A}$ a sus primeras dos componentes. Estas son "hojas" o subárboles binarios, $\mathcal{T}_0$ y $\mathcal{T}_1$.
   (b) Sean $f_0$ y $f_1$ las etiquetas de sus raices.

| $\mathcal{A}$ | [[8] [9]] | [5] | [6] | [2] | [3] | [0] | [1] | [4] | [7] |
|---|---|---|---|---|---|---|---|---|---|
| $f$ | 3 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

| $\mathcal{A}$ | [6] | [2] | [[[8] [9]] [5]] | [3] | [0] | [1] | [4] | [7] |
|---|---|---|---|---|---|---|---|---|
| $f$ | 4 | 5 | 6 | 6 | 7 | 8 | 9 | 10 |

| $\mathcal{A}$ | [[[8] [9]] [5]] | [3] | [0] | [1] | [[6] [2]] | [4] | [7] |
|---|---|---|---|---|---|---|---|
| $f$ | 6 | 6 | 7 | 8 | 9 | 9 | 10 |

| $\mathcal{A}$ | [0] | [1] | [[6] [2]] | [4] | [7] | [[[[8] [9]] [5]] [3]] |
|---|---|---|---|---|---|---|
| $f$ | 7 | 8 | 9 | 9 | 10 | 12 |

| $\mathcal{A}$ | [[6] [2]] | [4] | [7] | [[[[8] [9]] [5]] [3]] | [[0] [1]] |
|---|---|---|---|---|---|
| $f$ | 9 | 9 | 10 | 12 | 15 |

| $\mathcal{A}$ | [7] | [[[[8] [9]] [5]] [3]] | [[0] [1]] | [[[6] [2]] [4]] |
|---|---|---|---|---|
| $f$ | 10 | 12 | 15 | 18 |

| $\mathcal{A}$ | [[0] [1]] | [[[6] [2]] [4]] | [[7] [[[[8] [9]] [5]] [3]]] |
|---|---|---|---|
| $f$ | 15 | 18 | 22 |

| $\mathcal{A}$ | [[7] [[[[8] [9]] [5]] [3]]] | [[[0] [1]] [[[6] [2]] [4]]] |
|---|---|---|
| $f$ | 22 | 33 |

| $\mathcal{A}$ | [[[7] [[[[8] [9]] [5]] [3]]] [[[0] [1]] [[[6] [2]] [4]]]] |
|---|---|
| $f$ | 55 |

Recuadro 2: Computación del ejemplo.

(c) Sea $\mathcal{T} = \mathcal{T}_0 \triangle \mathcal{T}_1$ el árbol binario cuyo subárbol izquierdo es $\mathcal{T}_0$ y el derecho es $\mathcal{T}_1$. Etiquétese a su raíz con la suma $f := f_{01} = f_0 + f_1$, a su arista izquierda con 0 y a su arista derecha con 1.

(d) Insértese a $\mathcal{T}$ dentro de la lista $\mathcal{A}$ en el lugar que le corresponde según $f$, es decir, en la primera posición tal que todos los elementos de $\mathcal{A}$ hasta esa posición, tienen un peso estrictamente menor que $f$.

4. El código asociado a cada carácter es la lista dada por el camino desde la raíz hasta la hoja correspondiente al carácter.

Este algoritmo es de tipo voraz pues está codificando primeramente a los caracteres de mayor frecuencia.

**Ejemplo 3.1** *Consideremos el alfabeto Diez con las frecuencias siguientes:*

| *Carácter* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| *Frecuencia* $\cdot 10$ | *7* | *8* | *5* | *6* | *9* | *3* | *4* | *10* | *1* | *2* |

*Constrúyase el correspondiente Código de Huffman*

Puesto en orden no-decreciente, tenemos

| $\mathcal{A}$ | [8] | [9] | [5] | [6] | [2] | [3] | [0] | [1] | [4] | [7] |
|---|---|---|---|---|---|---|---|---|---|---|
| $f$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Procediendo según los pasos descritos en el algoritmo, de manera consecutiva, obtenemos la computación mostrada en la tabla (2). De lo cual resulta la codificación mostrada en la tabla (3). □

Así pues, si originalmente se hubiese ocupado $4 = \lceil \log_2(10) \rceil$ bits para representar cada símbolo de *Diez*, al sumar las frecuencias tendríamos que habría $\sum_{i=1}^{10} i = 55$ caracteres en el corpus original $\sigma$, y el "archivo"

$$
\begin{array}{lcl}
[7] & \rightsquigarrow & 00 \\
[4] & \rightsquigarrow & 111 \\
[1] & \rightsquigarrow & 101 \\
[0] & \rightsquigarrow & 100 \\
[3] & \rightsquigarrow & 011 \\
[2] & \rightsquigarrow & 1101 \\
[6] & \rightsquigarrow & 1100 \\
[5] & \rightsquigarrow & 0101 \\
[8] & \rightsquigarrow & 01000 \\
[9] & \rightsquigarrow & 01001
\end{array}
$$

Obsérvese que la codificación no es monótona respecto a las frecuencias y al orden lexicográfico. Sin embargo, esto no es una dificultad esencial, pues una leve modificación al algoritmo lograría la monotonía.

Recuadro 3: Codificación obtenida en el ejemplo.

que los contuviera tendría una longitud de 220 bits. Con la codificación de la tabla (3), el número total de bits será

$$10 \cdot 2 + (9 + 8 + 7 + 6) \cdot 3 + (5 + 4 + 3) \cdot 4 + (2 + 1) \cdot 5 = 173,$$

por lo que la razón de compresión será $\frac{173}{220} \approx 78.64\%$. □

## 3.2 Códigos de Huffman terciarios, cuaternarios y de orden mayor

Sea $k \geq 2$. Los *códigos de Huffman de orden k* se construyen sobre un alfabeto de $k$ símbolos, ordenados con un orden propio, y para ello se procede de igual manera a como se hizo en el caso binario. En cada iteración del ciclo principal se van agrupando de $k$ en $k$ vértices en el árbol que se construye.

## 3.3 Estimativo probabilista de la longitud de código

Dada una palabra $\sigma \in A^*$, la frecuencia $f(a) = \frac{c_\sigma(a)}{\text{long}(\sigma)}$ de un símbolo $a \in A$ puede identificarse como la *probabilidad de que ocurra a*. Si $\ell_a \in \mathbb{N}$ es la longitud del código de $a \in A$ entonces la *longitud esperada* del código de Huffman es

$$E(\ell) = \sum_{a \in A} \ell_a f(a) = \frac{1}{\text{long}(\sigma)} \sum_{a \in A} \ell_a c_\sigma(a).$$

Escribamos $p_a = f(a)$.

La *función de entropía* $H : \mathbb{R}^* \to \mathbb{R}$ se define de manera que cumpla con las propiedades siguientes:

**Positiva.** $\forall p_1, \ldots, p_n$: $H(p_1, \ldots, p_n) \geq 0$.

**Contínua.** $(p_1, \ldots, p_n) \mapsto H(p_1, \ldots, p_n)$ es contínua.

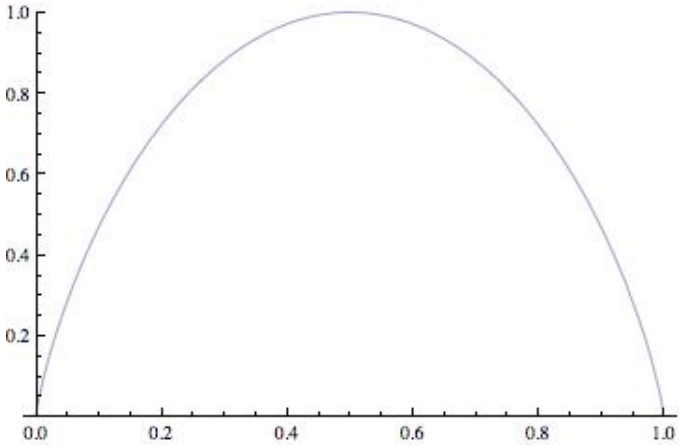
**Simétrica.** $\forall p_1, \ldots, p_n, \forall \pi \in S_n : H(p_1, \ldots, p_n) = H\left(p_{\pi(1)}, \ldots, p_{\pi(n)}\right)$.

**Coherente.** $\forall p_1, p_2, \ldots, p_n$: $H(p_1, p_2, \ldots, p_n) = H(p_1 + p_2, \ldots, p_n) + (p_1 + p_2) H\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}, \ldots, p_n\right)$.

Para esto se debe tener que existe $C > 0$ tal que

$$\forall p_1, \ldots, p_n : \quad H(p_1, \ldots, p_n) = C \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$$

Dos símbolos




Tres símbolos

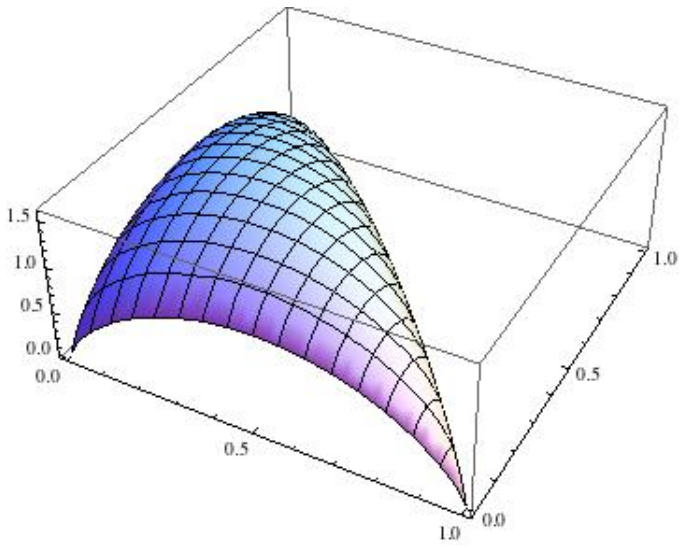



Figura 1: Gráfica de la función de entropía de Shannon.

donde se ha de entender: $\left[p = 0 \;\Rightarrow\; p \log \frac{1}{p} = 0.\right]$. Al expresar $C = (\log c)^{-1}$, se ha de tener

$$H\left(p_1, \ldots, p_n\right) = \sum_{i=1}^{n} p_i \log_c \frac{1}{p_i} = -\sum_{i=1}^{n} p_i \log_c p_i. \tag{6}$$

Esta última expresión define la llamada *entropía de Shannon*, considerando $c = 2$. Para el caso $n = 2$, la relación (6) queda

$$\forall p \in [0,1]: \; H(p, 1-p) = -p \log_2 p - (1-p) \log_2 (1-p).$$

Esta función es tal que $H([0,1]) = [0,1]$, posee sus valores mínimos en $p = 0, 1$ (que una frecuencia sea 1 significa que en la cadena original sólo ha de aparecer un símbolo) y su valor máximo en $p = \frac{1}{2}$ (los símbolos que aparecen en la cadena original son equiprobables).

Sea $\mathbf{e}_i = (\delta_{ij})_{j=1}^{n}$ el $i$-ésimo vector canónico que tiene el valor 1 en la $i$-ésima coordenada y el valor 0 en las otras. Este corresponde a la distribución de probabilidad en la que sólo aparece el $i$-ésimo símbolo y ninguno otro. Pues bien, de la relación (6) se ve que $H(\mathbf{e}_i) = 0$, lo que da una mínima entropía. Si se considera $\mathbf{f} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{e}_i$, que corresponde a la distribución uniforme de probabilidad, se tiene $H(\mathbf{f}) = \log_2(n)$, lo que da una máxima entropía.

En la figura 1 presentamos las gráficas de $H$ para alfabetos de dos y de tres símbolos respectivamente.

Supongamos que $\mathbf{p} = (p_j)_{j=1}^n$ es la distribución de probabilidad de un alfabeto $A$ de $n$ símbolos, y que $\mathbf{l} = (\ell_j)_{j=1}^n$ es la lista de longitudes de un código instantáneo de $A$. Entonces la longitud esperada ha de ser $E(\mathbf{l}) = \sum_{j=1}^n \ell_j p_j = \sum_{j=1}^n p_j \log_2 \left[2^{\ell_j}\right]$ y, en consecuencia

$$H(\mathbf{p}) - E(\mathbf{l}) = \sum_{j=1}^n p_j \log_2 \left[\frac{1}{p_j}\right] - \sum_{j=1}^n p_j \log_2 \left[2^{\ell_j}\right] = \sum_{j=1}^n p_j \log_2 \left[\frac{1}{2^{\ell_j} p_j}\right] = \frac{1}{\log 2} \sum_{j=1}^n p_j \log \left[\frac{1}{2^{\ell_j} p_j}\right]$$

de donde

$$H(\mathbf{p}) - E(\mathbf{l}) \leq \frac{1}{\log 2} \sum_{j=1}^n p_j \left[\frac{1}{2^{\ell_j} p_j} - 1\right] = \frac{1}{\log 2} \left[\sum_{j=1}^n \frac{1}{2^{\ell_j}} - 1\right] = \frac{1}{\log 2} \left[\sum_{j=1}^n 2^{-\ell_j} - 1\right] \leq 0$$

(donde la última es precisamente la desigualdad de Kraft). Así pues, $H(\mathbf{p}) \leq E(\mathbf{l})$, es decir *la longitud esperada de cualquier código instantáneo es mayor o igual que la entropía de la distribución de probabilidad de los símbolos.*

Para una distribución $\mathbf{p} = (p_j)_{j=1}^n$ de un alfabeto $A$ de $n$ símbolos y un entero positivo $k \in N$, se define la distribución $\mathbf{p}^k$ sobre el alfabeto $A^k$, que consiste de las palabras de longitud $k$, haciendo

$$\forall \mathbf{a} = a_{i_0} \cdots a_{i_{k-1}} \in A^k : \quad p_{i_0 \cdots i_{k-1}} = \prod_{\kappa=1}^k p_{i_\kappa}.$$

$\mathbf{p}^k$ se dice ser la *k-ésima extensión de* $\mathbf{p}$.

Se tendrá entonces que vale el

**Teorema 3.1 (de Shannon de Códigos sin Ruido)** *Para cualquier distribución de probabilidad* $\mathbf{p}$*, sea* $\ell_H(\mathbf{p})$ *la longitud esperada de una codificación de Huffman. Entonces*

$$H(\mathbf{p}) \leq \ell_H(\mathbf{p}) \leq H(\mathbf{p}) + 1.$$

*Y para extensiones sucesivas,*

$$\lim_{k \to +\infty} \frac{1}{k} \ell_H(\mathbf{p}^k) = H(\mathbf{p}).$$

Planteemos un caso de estudio que se resolvería de manera directa utilizando el teorema de Shannon 3.1.

**Ejemplo 3.2** *Supongamos que se quiere transmitir un croquis. Este está conformado prácticamente por unas cuantas líneas sobre fondo blanco. Al digitalizar la imagen, la probabilidad de que aparezca un* 1 *(pixel negro) es a lo sumo* $p$ *y de que aparezca* 0 *(pixel blanco) es al menos* $1 - p$*, con* $p < 10^{-2}$*, digamos. El alfabeto original* $A = \{0, 1\}$ *consta de dos símbolos. Para* $k = 10$*, se divide la imagen en bloques de* $k$ *pixeles consecutivos. El alfabeto actual es entonces* $A^k$*. A toda cadena de ceros,* $0^{(k)}$*, se la codifica con un solo* 0*, y a cualquier otra cadena* $\sigma \in A^k$ *con la cadena* $1\sigma \in A^{k+1}$*. ¿Cuál es la esperanza de la longitud del código?*

La longitud esperada del código de una cadena $\tau \in A^k$ es

$$\begin{aligned} E(\ell) &= 1 \cdot \text{Prob}\left(\tau = 0^{(k)}\right) + (1 + k)\,\text{Prob}\left(\tau \neq 0^{(k)}\right) \\ &= 1 \cdot (1 - p)^k + (1 + k)(1 - (1 - p)^k) \\ &= 1 + k - k(1 - p)^k, \end{aligned}$$

así esta longitud será más cercana a 1 conforme $p$ sea más pequeño, o sea la compresión será mayor (de $k$ a 1). El teorema de Shannon 3.1 da la misma estimación de manera más directa y general.

## 3.4 Primeras listas

### 3.4.1 Ejercicios

1. Diseñe un código instantáneo para un alfabeto llano de $n$ símbolos, de manera que las longitudes de cada uno de los códigos de los símbolos coincidan. Indique cuál es el número $k$, en términos de $n$, de símbolos necesarios en el alfabeto código.

2. Decida si el código definido por la siguiente tabla es de decodificación única:

| | | | | | |
|---|---|---|---|---|---|
| 1 | 01 | 3 | 10 | 5 | 1100 |
| 2 | 011 | 4 | 1000 | 6 | 0111 |

Si no lo fuera, localice dos palabras distintas que produzcan un mismo código.

3. Decida si el código definido por la siguiente tabla es de decodificación única:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | aa | 2 | abbbb | 4 | abbaa | 6 | bbbab | 8 | aaaaba |
| 1 | aabab | 3 | ababa | 5 | babba | 7 | aaaabb | 9 | aaaaab |

Si no lo fuera, localice dos palabras distintas que produzcan un mismo código.

4. Sea $A = (a_\nu)_{\nu=0}^{n-1}$ un alfabeto de $n$ símbolos. Sea $\sigma \in A^*$ una palabra tal que, para cada $\nu$ tal que $1 \leq \nu \leq n-1$ el símbolo $a_\nu$ aparece el doble de veces que $a_{\nu-1}$. Describa cómo ha de ser el código de Huffman binario.

5. Sea $A = (a_\nu)_{\nu=0}^{n-1}$ un alfabeto de $n$ símbolos. Sea $\sigma \in A^*$ una palabra tal que todos los símbolos aparecen un mismo número de veces en $\sigma$. Describa cómo ha de ser el código de Huffman binario. Trate por separado el caso en que $\text{long}(\sigma)$ es una potencia de 2 y el caso complementario.

6. Dé un ejemplo de una tabla de frecuencias para un alfabeto de 5 símbolos de manera que el código binario de Huffman tenga una longitud promedio 1.8.

7. En un mensaje $\sigma$ sobre un alfabeto de 4 símbolos, $A = \{a_0, a_1, a_2, a_3\}$, se tiene que el primero, $a_0$ aparece siete veces más que cualquiera de los otros. Construya un código binario que tenga una longitud promedio 1.4.

8. Suponga que se tiene un alfabeto de 256 símbolos de ocurrencias equiprobables. ¿Cuál es el valor de su entropía? ¿Cuántos deben ser los símbolos en el alfabeto para que en ocurrencias equiprobables se tenga una entropía de 50 bits?

9. La *efectividad* $\Phi : \mathbf{p} \mapsto \Phi(\mathbf{p})$ de una distribución $\mathbf{p}$ se define como la razón entre la entropía $H(\mathbf{p})$ y la longitud promedio del código binario de Huffman $\ell_H(\mathbf{p})$ que determina. Encuentre los valores extremos de la efectividad e indique a cuáles distribuciones corresponden.

10. Aplique el teorema de Shannon de Códigos sin Ruido 3.1 a la situación descrita en el ejemplo 3.2 y calcule estimativos, en términos de $k$ y de $p$, de la tasa de compresión esperada.

### 3.4.2 Programas

1. **Enumeración de subconjuntos.** Escriba un programa que reciba dos enteros positivos $(n, k)$, con $0 < k \leq n$ y haga lo siguiente:

   A. Reciba un índice $\ell \in [\![0, \binom{n}{k} - 1]\!]$ y escriba el $\ell$-ésimo subconjunto de $k$ elementos en $[\![0, n-1]\!]$.

   B. Reciba un subconjunto $I$ de $k$ elementos en $[\![0, n-1]\!]$ y calcule el índice que le corresponda de acuerdo con la enumeración anterior.

2. **Código 4-en-8.** Sea $A$ el alfabeto de 70 símbolos

$$A = \{\texttt{A,...,Z}\} \cup \{\texttt{a,...,z}\} \cup \{\texttt{0,...,9}\} \cup \{\texttt{!,?,(,),:,., ,}\}$$

(en el último bloque el penúltimo carácter es el *blanco* y el último la *coma*). Escriba un programa que considerando el Código 4-en-8 $\gamma$ del ejemplo 2.1:

A. Reciba una palabra $\sigma \in A^*$ y escriba el código $\gamma(\sigma)$ correspondiente a esa palabra.

B. Reciba una palabra $\omega \in \{\texttt{0},\texttt{1}\}^*$ y revise si acaso está en la imagen de $\gamma$. Si no lo está marque el primer lugar en $\omega$ donde ocurra una violación a las reglas sintácticas, y si lo está entonces encuentre $\sigma \in A^*$ tal que $\gamma(\sigma) = \omega$.

3. Escriba un programa que habiendo recibido como entrada una lista de números enteros positivos $(\ell_\nu)_{\nu=0}^{n-1}$:

A. Calcule el mínimo $k \in \mathbb{N}$ que satisface la desigualdad de Kraft,

B. Sea $A_n$ el alfabeto consistente de los $n$ símbolos consecutivos ASCII a partir del carácter numérico con código decimal 33 ("!"). Sea $C_k$ el alfabeto consistente de los $k$ símbolos consecutivos del alfabeto latino de minúsculas. Calcule la función de codificación de un código instantáneo de $A_n$ sobre $C_k^*$ tal que el símbolo $i$-ésimo se codifica por una cadena de longitud $\ell_i$.

4. Escriba un programa que reciba como entrada la tabla correspondiente a la función de codificación de un código.

A. Decida si el código es instantáneo,

B. En caso de que lo sea, reciba palabras codificadas y las decodifique procediendo de manera voraz..

5. Escriba un programa que habiendo recibido como entrada una lista de números enteros positivos $(\ell_\nu)_{\nu=0}^{n-1}$ y un entero $k \in \mathbb{N}$ que satisfagan la desigualdad de Kraft, cuente cuántas funciones de codificación existen tales que el símbolo $i$-ésimo se codifica por una cadena de longitud $\ell_i$.

6. Escriba un programa que calcule la codificación de Huffman. Como entrada debe darse una tabla de símbolos y frecuencias y una bandera que indique si se quiere un código binario, terciario o cuternario. Como salida debe dar la función de codificación correspondiente.

7. Escriba un programa que reciba un entero positivo $n \in \mathbb{N}$ y un valor real $h \in \mathbb{R}$. Para ellos debe decidir si existe una distribución $\mathbf{p}$ tal que $H(\mathbf{p}) = h$ y en tal caso debe encontrar una tal distribución, la longitud esperada de su código de Huffman y su efectividad.

8. Escriba un programa que reciba una distribución $\mathbf{p}$ de un alfabeto de $n$ símbolos, y un entero $k \in \mathbb{N}$ y calcule la extensión $\mathbf{p}^k$ como una lista de $n^k$ valores reales.

9. Escriba un programa que reciba como entrada $n \in \mathbb{N}$ y un vector $\mathbf{p} \in \mathbb{R}^n$ y compruebe experimentalmente que vale la igualdad de límite enunciada en el teorema de Shannon 3.1.

10. Escriba un programa que realice la situación descrita en el ejemplo 3.2:

- Inicialmente recibe $n, m \in \mathbb{N}$ y $p \in [0, 1]$.
- Genera una matriz $M \in \{0, 1\}^{m \times n}$ de manera que la frecuencia de 1 sea $p$ (y en consecuencia la de 0 habría de ser $1 - p$).
- Realiza la codificación mostrada en el ejemplo, y calcula la tasa de compresión.
- Muestra estadísticas para muchas repeticiones del experimento.

# 4 Códigos binarios

Sea $\mathbb{F}_2 = \{0, 1\}$ el campo primo consistente de dos elementos con sus operaciones de suma y producto usuales (XOR y AND respectivamente). Las potencias cartesianas $\mathbb{F}_2^n$ tienen naturalmente una estructura de espacio vectorial sobre $\mathbb{F}_2$. Un *código binario* es uno de la forma $\gamma : \mathbb{F}_2 \to \mathbb{F}_2^*$.

## 4.1 Código por mayoría de votos

Supongamos que se quiere transmitir mensajes, desde una parte *emisora* a una *receptora*, por un canal que comete errores con una distribución binomial, es decir, existe un $p \in [0,1]$ tal que para cada $x \in \mathbb{F}_2$, si $\tilde{x}$ es el valor recibido a través del canal, entonces la *probabilidad de error* es

$$\text{Prob}\{\tilde{x} \neq x\} = \text{Prob}\{\tilde{x} + x = 1\} = p.$$

Como un primer paso para detectar errores se podría utilizar, para un número $n \in \mathbb{N}$ impar, el código $\kappa_n : x \mapsto x^{(n)}$, que codifica cada bit $x$ mediante $n$ repeticiones consecutivas de él mismo. En la parte del receptor, para decodificar un bloque $\sigma$ de $n$ bits, se procede por *simple mayoría*: el valor $x$ que aparezca más veces en $\sigma$ es el que se toma como transmitido.

La probabilidad de cometer un error en la decodificación es la probabilidad de que habiendo transmitido un bit $x$, su valor complementario sea mayoritario en $\sigma$. Así pues ésta es:

$$P_{en} = \sum_{j=\frac{n+1}{2}}^{n} \binom{n}{j} p^j (1-p)^{n-j} = \sum_{i=0}^{\frac{n-1}{2}} \binom{n}{i} p^{n-i} (1-p)^i.$$

Para $p < \frac{1}{2}$ se tiene $P_{en} \to 0$ conforme $n$ aumenta de tamaño; para $p = \frac{1}{2}$ se tiene $P_{en} = \frac{1}{2}$ independientemente del valor de $n$; pero para $p > \frac{1}{2}$ se tiene $P_{en} \to 1$ conforme $n$ aumenta de tamaño.

## 4.2 Distancia de Hamming

Para $n \in \mathbb{N}$ hay $2^n$ palabras de longitud $n$ con símbolos 0, 1. Sea $k \in [\![1, n]\!]$ y sea $K \subset [\![0, n-1]\!]$ un conjunto de índices de cardinalidad $k$. Sea $C \subset \mathbb{F}_2^n$ un conjunto de cardinalidad $2^k$ y sea $\gamma : \mathbb{F}_2^k \to C$ tal que existe $\iota : [\![0, k-1]\!] \to K$ biyectiva, con la condición:

$$\forall \boldsymbol{\varepsilon} \in \mathbb{F}_2^k \ \forall i \in [\![0, k-1]\!] : \ (\boldsymbol{\varepsilon})_i = (\gamma(\boldsymbol{\varepsilon}))_{\iota(i)}.$$

Se dice que el código $\gamma$ *posee $k$ bits de información* y *utiliza $n-k$ bits de revisión*. El cociente $R_C = \frac{k}{n}$ es la *razón de información* del código $C$.

Por ejemplo, para el código por mayoría de votos visto en la sección anterior 4.1, $n$ es un número impar, $k = 1$ y, por ejemplo, $K = \{0\}$. $C = \{0^n, 1^n\}$ y $\gamma : \varepsilon \mapsto \varepsilon^n$. Entonces ese código tiene un solo bit de información, utiliza $n-1$ bits de revisión y su razón de información es $\frac{1}{n}$.

**Ejemplo 4.1** *Sea $C = \{\boldsymbol{\varepsilon} \in \mathbb{F}_2^n |\ \varepsilon_{n-1} = \left(\sum_{i=0}^{n-2} \varepsilon_i\right) \bmod 2\}$ el conjunto de palabras cuyo último símbolo es la paridad del bloque formado por los primeros $n-1$. Para $k = n-1$ y $K = [\![0, n-2]\!]$, las funciones $\gamma$ e $\iota$ quedan determinadas naturalmente: $\gamma : \boldsymbol{\varepsilon} \mapsto (\boldsymbol{\varepsilon}, \sum_{i=0}^{n-2} \varepsilon_i)$, $\iota : i \mapsto i$. Entonces $\gamma$ es un código con $n-1$ bits de información y $1 = n - (n-1)$ bit de revisión. La razón de información de este código es $\frac{n-1}{n}$.*

**Ejemplo 4.2** *Sean $n = 6$ y $k = 3$. Definamos*

$$\gamma : \begin{array}{ccc} 000 & \mapsto & 000000 \\ 001 & \mapsto & 001110 \\ 010 & \mapsto & 010101 \\ 011 & \mapsto & 011011 \\ 100 & \mapsto & 100011 \\ 101 & \mapsto & 101101 \\ 110 & \mapsto & 110110 \\ 111 & \mapsto & 111000 \end{array}$$

*Sea $K = [\![0, 2]\!] \subset [\![0, 5]\!]$ e $\iota : [\![0, 2]\!] \to K$ la identidad. $\gamma$ es un código con 3 bits de información y $3 = 6 - 3$ bits de revisión, y su razón de información es $\frac{1}{2}$.*

Observemos que en el ejemplo anterior, cualesquiera dos palabras en el código $C$ difieren en por lo menos 3 bits. Así pues, habiendo recibido una palabra $\boldsymbol{\varepsilon} \in \mathbb{F}_2^6$, si ésta no está en $C$ se reconoce que hay un error.

Las palabras con error pueden ser complementarias de palabras de código o las restantes (éstas son $2^6 - 2 \cdot 8 = 2^6 - 2^4 = 2^4 \cdot 3 = 48$) y para cada una de estas últimas habrá una única palabra $\boldsymbol{\varepsilon}' \in C$ que difiera de $\boldsymbol{\varepsilon}$ en tan solo un bit. Se toma entonces a la palabra correspondiente con $\boldsymbol{\varepsilon}'$ según $\gamma$ y de esta manera se corrige el error. (Para cada palabra complementaria de una palabra de código hay tres palabras de código que difieren de ella en dos bits.)

**Definición 4.1 (Distancia de Hamming)** *Sea* $d_n : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{N}$,

$$(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1) \mapsto d_n(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1) = \text{card}\{j \in [\![0, n-1]\!] |\ \varepsilon_{0j} \neq \varepsilon_{1j}\},$$

*la función que a cada par de palabras le asocia el número de sus discordancias. Naturalmente,* $d_n$ *es una función distancia en* $\mathbb{F}_2^n$*, llamada* de Hamming.

**Definición 4.2** *Para cada* $\boldsymbol{\varepsilon} \in \mathbb{F}_2^n$ *y* $r \in \mathbb{N}$ *la* esfera *de* radio $r$ centrada en $\boldsymbol{\varepsilon}$ *es*

$$S(\boldsymbol{\varepsilon}, r) = \{\boldsymbol{\delta} \in \mathbb{F}_2^n |\, d_n(\boldsymbol{\varepsilon}, , \boldsymbol{\delta}) = r\},$$

*y la* bola *de* radio $r$ centrada en $\boldsymbol{\varepsilon}$ *es*

$$B(\boldsymbol{\varepsilon}, r) = \{\boldsymbol{\delta} \in \mathbb{F}_2^n |\, d_n(\boldsymbol{\varepsilon}, \boldsymbol{\delta}) \leq r\} = \bigcup_{s \in [\![0,r]\!]} S(\boldsymbol{\varepsilon}, s).$$

Así, se tiene que $\text{card}\,(S(\boldsymbol{\varepsilon}, r)) = \binom{n}{r}$ y $\text{card}\,(B(\boldsymbol{\varepsilon}, r)) = \sum_{s \in [\![0,r]\!]} \binom{n}{s}$.

**Definición 4.3** *Para un punto* $\boldsymbol{\varepsilon} \in \mathbb{F}_2^n$ *y un conjunto* $G \subset \mathbb{F}_2^n$ *se define* $d_n(\boldsymbol{\varepsilon}, G) = \min\{d_n(\boldsymbol{\varepsilon}, \boldsymbol{\delta}) |\ \boldsymbol{\delta} \in G\}$. *Para dos conjuntos* $G_0, G_1 \subset \mathbb{F}_2^n$ *se define* $d_n(G_0, G_1) = \min\{d_n(\boldsymbol{\varepsilon}, G_1) |\ \boldsymbol{\varepsilon} \in G_0\}$.

**Definición 4.4** *Para un conjunto* $G \subset \mathbb{F}_2^n$ *su* diámetro *es*

$$d_{n,max}(G) = \max\left\{d_n\left(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1\right) |\ \boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in G\right\}.$$

*Su* distancia mínima *es*

$$d_{n,min}(G) = \min\left\{d_n\left(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1\right) |\ \boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1 \in C \ \&\ \boldsymbol{\varepsilon}_0 \neq \boldsymbol{\varepsilon}_1\right\}.$$

**Observación 4.1** *Para cualquier código* $C \subset \mathbb{F}_2^n$ *tal que* $\left(B(\boldsymbol{\varepsilon}, d_{n,min}(C) - 1)\right)_{\boldsymbol{\varepsilon} \in C}$ *sea un recubrimiento de* $\mathbb{F}_2^n$ *se tiene*

- $C$ *reconoce* $t$ *errores si y sólo si* $d_{n,min}(C) > t$
- $C$ *corrige* $t$ *errores si y sólo si* $d_{n,min}(C) > 2t$

En efecto, para la primera aseveración: Supóngase que se quiere enviar la palabra $\boldsymbol{\varepsilon} \in C$ y se recibe $\boldsymbol{\delta} \in \mathbb{F}_2^n$. Entonces pueden ocurrir dos casos. Si $\boldsymbol{\delta} \notin C$, se habrá detectado un error y como éste está en una bola de radio $d_{n,min}(C) - 1$ con centro en un punto de $C$ se ve que hay menos de $d_{n,min}(C)$ errores. Si, por lo contrario, $\boldsymbol{\delta} \in C$, entonces bien $\boldsymbol{\delta} = \boldsymbol{\varepsilon}$, en cuyo caso no hay error, o bien $d_n(\boldsymbol{\delta}, \boldsymbol{\varepsilon}) \geq d_{n,min}(C)$ y no será posible detectar los más de $t$ errores cometidos.

La segunda aseveración se sigue de lo siguiente: Supóngase que se quiere enviar la palabra $\boldsymbol{\varepsilon}_0 \in C$ y se recibe $\boldsymbol{\delta} \in \mathbb{F}_2^n$ con a lo sumo $t$ errores. Entonces $d_n(\boldsymbol{\delta}, \boldsymbol{\varepsilon}_0) \leq t$. Sea $\boldsymbol{\varepsilon}_1 \in C$ una palabra en el código, distinta de la enviada. Entonces $d_n(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1) \geq d_{n,min}(C)$. Resulta pues

$$d_n(\boldsymbol{\delta}, \boldsymbol{\varepsilon}_0) \geq d_n(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1) - d_n(\boldsymbol{\delta}, \boldsymbol{\varepsilon}_0) > d_{n,min}(C) - t.$$

Por tanto: $d_n(\boldsymbol{\delta}, \boldsymbol{\varepsilon}_0) > t \ \Leftrightarrow\ d_{n,min}(C) > 2t$. □

**Definición 4.5** *Un código $C \subset \mathbb{F}_2^n$ con $k$ bits de información y $n-k$ bits de revisión se dice ser un* código-$[n,k]$. *En este caso se dice también que $k$ es la* dimensión *de $C$ y que $n$ es su* longitud. *Todo código-$[n,k]$ de distancia mínima $d$ se dice ser un* código-$[n,k,d]$.

La observación 4.1 se reformula como la siguiente:

**Observación 4.2** *Todo código-$[n,k,d]$ puede corregir menos de $\left\lfloor \frac{d}{2} \right\rfloor$ errores.*

**Ejemplo 4.3** *Sean $n=5$ y $k=2$. Definamos*

$$\gamma : \begin{array}{ccc} 00 & \mapsto & 00000 \\ 01 & \mapsto & 00111 \\ 10 & \mapsto & 11100 \\ 11 & \mapsto & 11011 \end{array}$$

*Sea $K = \{0,3\} \subset [\![0,4]\!]$ e $\iota : [\![0,1]\!] \to K$, $i \mapsto 3i$. Se tiene que $\gamma$ es un código-$[5,2]$ y su razón de información es $\frac{2}{5}$. Su distancia mínima es 3, por lo que es un código-$[5,2,3]$*

Por la observación 4.2, se tiene que puede corregir un solo bit. Alrededor de cada palabra de código, la bola de radio 1 consiste de $\binom{5}{0} + \binom{5}{1} = 1 + 5 = 6$ elementos. Se tiene;

$$\begin{aligned} B(00000,\,1) &= \{00000, 00001, 00010, 00100, 01000, 10000\} \\ B(00111,\,1) &= \{00111, 00110, 00101, 00011, 01111, 10111\} \\ B(11100,\,1) &= \{11100, 11101, 11110, 11000, 10100, 01100\} \\ B(11011,\,1) &= \{11011, 11010, 11001, 11111, 10011, 01011\} \end{aligned}$$

Estas bolas son ajenas a pares y su unión consiste de 24 palabras. Las restantes $8 = 2^5 - 24$ equidistan en 2 de dos palabras de código. □

## 4.3 Códigos lineales: Códigos rectangulares y de Hamming

**Definición 4.6** *Un código $C \subset \mathbb{F}_2^n$ se dice ser* lineal *si posee al origen y es cerrado bajo la suma de $\mathbb{F}_2^n$.*

*En el caso de un código lineal, el* peso *de cualquier palabra $\boldsymbol{\varepsilon} \in C$ es $w(\boldsymbol{\varepsilon}) = d_n(\boldsymbol{\varepsilon}, \mathbf{0})$, y el* peso mínimo *del código es*

$$w_m(C) = \min\{w(\boldsymbol{\varepsilon}) \mid \boldsymbol{\varepsilon} \in C - \{\mathbf{0}\}\}.$$

**Observación 4.3** *Para cualquier código lineal $C \subset \mathbb{F}_2^n$ se tiene*

- *$C$ reconoce $t$ errores si $w_m(C) > t$*
- *$C$ corrige $t$ errores si $w_m(C) > 2t$*

**Ejemplo 4.4 (Códigos rectangulares)** *Supongamos $n_1 = m\,n$, con $m, n > 1$. Entonces el producto cartesiano $[\![0, m-1]\!] \times [\![0, n-1]\!]$ se identifica con el intervalo de enteros $[\![0, n_1 - 1]\!]$ mediante la correspondencia $\nu : (i,j) \mapsto \nu(i,j) = in + j$. El* código rectangular*, con $(m-1)(n-1)$ bits de información y $m+n-1$ bits de revisión, es*

$$\begin{aligned} C \;=\; \Bigg\{ \boldsymbol{\varepsilon} \in \mathbb{F}_2^{n_1} \mid \; & \left[ \forall i \in [\![0, m-1]\!] :\; \varepsilon_{\nu(i,n-1)} = \left( \sum_{j=0}^{n-2} \varepsilon_{\nu(i,j)} \right) \bmod 2 \right] \;\&\; \\ & \left[ \forall j \in [\![0, n-1]\!] :\; \varepsilon_{\nu(m-1,j)} = \left( \sum_{i=0}^{m-2} \varepsilon_{\nu(i,j)} \right) \bmod 2 \right] \Bigg\} \end{aligned}$$

*La razón de información de este código es $\frac{m-1}{m}\frac{n-1}{n}$.*

Puede verse que $C$ es un código lineal, pues contiene al origen y es cerrado bajo la suma. Si un elemento $\varepsilon_{\nu(i,j)}$ toma el valor 1, entonces debe aparecer un 1 en otra posición en la misma columna y un 1 también en otra posición en el mismo renglón, y debe haber un 1 en la esquina opuesta a $(i,j)$. En consecuencia, el peso mínimo de $C$ es 4 y por tanto es capaz de detectar hasta 3 errores.

Se tiene que un vector $\boldsymbol{\varepsilon} \in \mathbb{F}_2^{n_1}$ estará en el código si y sólo si se satisfacen las $m+n$ ecuaciones

$$\sum_{j=0}^{n-1} \varepsilon_{\nu(i,j)} = 0,\ i \in [\![0, m-1]\!] \ \ \& \ \ \sum_{i=0}^{m-1} \varepsilon_{\nu(i,j)} = 0,\ j \in [\![0, n-1]\!].$$

Todo código lineal es un subespacio lineal de $\mathbb{F}_2^n$ y por tanto ha de poseer una base. Cualquier palabra podrá ser escrita como una combinación lineal de los elementos en la base y por tanto las palabras en el código han de satisfacer un conjunto de ecuaciones lineales tal como en el caso de los códigos rectangulares.

En efecto, si un código lineal $C$ es de dimensión $k$ en $\mathbb{F}_2^n$, sea $(\boldsymbol{\varepsilon}_j)_{0\le j\le k} \subset \mathbb{F}_2^n$ una base de $C$. La matriz $H = [\boldsymbol{\varepsilon}_j]_{0\le j\le k} \in \mathbb{F}_2^{n\times k}$ que posee como columnas a los vectores básicos, llamada *generatriz* de $C$, es de orden $(n\times k)$ y determina un isomorfismo $\mathbb{F}_2^k \to C$, $\boldsymbol{\delta} \mapsto H\boldsymbol{\delta}$. Existe entonces una matriz $H^\perp \in \mathbb{F}_2^{(n-k)\times n}$ tal que $H^\perp H = 0 \in \mathbb{F}_2^{(n-k)\times k}$. Así se ha de tener la equivalencia: $\left[\boldsymbol{\varepsilon} \in C \ \Leftrightarrow\ H^\perp\boldsymbol{\varepsilon} = \mathbf{0}\right]$. La matriz $H^\perp$ se dice ser una *revisora de paridad* del código $C$.

**Proposición 4.1** *Un código $C$ corrige errores de un bit si y sólo si cualquier matriz revisora de paridad suya posee columnas no-nulas y distintas a pares.*

Sea $\mathbf{e}_j = [\delta_{ij}]_{i=0}^{n-1}$ el $j$-ésimo vector de la base canónica, $j = 0,\dots,n-1$. Denotemos también por $\boldsymbol{\varepsilon}_{ij}$ al vector que coincide con $\mathbf{0}$ salvo que en sus dos entradas $i$ y $j$ posee el valor 1, $\boldsymbol{\varepsilon}_{ij} = \mathbf{e}_i + \mathbf{e}_j$.

La proposición se sigue de que las siguientes aseveraciones son equivalentes a pares:

- $C$ corrige errores de un bit.
- El peso mínimo de $C$ es al menos 3.
- Ninguno de los vectores $\mathbf{e}_j$, $\boldsymbol{\varepsilon}_{ij}$ puede estar en $C$.
- Para cada $H^\perp$ revisora de paridad, los productos $H^\perp\mathbf{e}_j$, $H^\perp\boldsymbol{\varepsilon}_{ij}$ no pueden ser nulos.
- Para cada $H^\perp$ revisora de paridad, las columnas de $H^\perp$ son no-nulas y distintas a pares.

**Definición 4.7** *Para cada $m$, sea $H_m^\perp = [\boldsymbol{\varepsilon}]_{\boldsymbol{\varepsilon}\in\mathbb{F}_2^m - \{\mathbf{0}\}} \in \mathbb{F}_2^{m\times(2^m-1)}$ la matriz cuyas columnas son los vectores no-nulos en $\mathbb{F}_2^m - \{\mathbf{0}\}$. Todo código que posea a $H_m^\perp$ como matriz revisora de paridad se dice ser* de Hamming[1].

Así todo código de Hamming posee $(2^m-1)-m$ bits de información y $m$ bits de revisión, y su razón de información es $\frac{2^m-1-m}{2^m-1} = 1 - \frac{m}{2^m-1}$.

La matriz $H_m^\perp$ queda determinada de manera única salvo el orden en el que se enumere a los elementos de $\mathbb{F}_2^m - \{\mathbf{0}\}$. Sea $\kappa_m : [\![1, 2^m-1]\!] \to [\![1, 2^m-1]\!]$ la permutación que ordena a los índices de acuerdo con el peso de Hamming y en orden lexicográfico cuando haya coincidencia de pesos. Por ejemplo:

$m = 3.\ \kappa_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 4 & 3 & 5 & 6 & 7 \end{pmatrix}$

$m = 4.\ \kappa_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 2 & 4 & 8 & 3 & 5 & 6 & 9 & 10 & 12 & 7 & 11 & 13 & 14 & 15 \end{pmatrix}$

Si a los índices se les ordena de acuerdo con $\kappa_m$, entonces se podrá escribir $H_m^\perp = [I_m\ G_m]$ donde $I_m$ es la matriz identidad de orden $(m\times m)$ y $G_m$ es una matriz de orden $m\times(2^m-1-m)$. Así pues, se tendrá que una palabra $\boldsymbol{\varepsilon} = (\varepsilon_j)_{j=1}^{2^m-1}$ está en el código de Hamming si y sólo si $[I_m\ G_m]\ \kappa_m(\boldsymbol{\varepsilon}) = \mathbf{0}$, donde $\kappa_m(\boldsymbol{\varepsilon}) = \left(\varepsilon_{\kappa_m(j)}\right)_{j=1}^{2^m-1}$, lo cual equivale a que

$$I_m\boldsymbol{\varepsilon}_0 = G_m\boldsymbol{\varepsilon}_1 \text{ , donde } \boldsymbol{\varepsilon}_0 = \left(\varepsilon_{\kappa_m(j)}\right)_{j=1}^{m} \text{ y } \boldsymbol{\varepsilon}_1 = \left(\varepsilon_{\kappa_m(j)}\right)_{j=m+1}^{2^m-1}. \tag{7}$$

[1] Estos fueron presentados en 1947 por Richard Hamming de los Laboratorios Bell en los EEUU.

El conjunto de índices $I_0 = \{\kappa_m(j)\}_{j=1}^{m}$ corresponde a los bits de revisión y el conjunto $I_1 = \{\kappa_m(j)\}_{j=m+1}^{2^m-1}$ a los de información.

Sea $H_m = \begin{bmatrix} G_m \\ I_{2^m-1-m} \end{bmatrix} \in \mathbb{F}_2^{(2^m-1)\times(2^m-1-m)}$. Entonces, $H_m^{\perp} H_m = \mathbf{0} \in \mathbb{F}_2^{m\times(2^m-1-m)}$ y por tanto las columnas de $H_m$ forman una base del código de Hamming. La dimensión del código es el número de bits de información, a saber, $2^m - 1 - m$.

Para codificar una palabra $\delta = [\delta_j]_{j=1}^{2^m-1-m} \in \mathbb{F}_2^{2^m-1-m}$ se construye $\boldsymbol{\varepsilon} = (\varepsilon_j)_{j=1}^{2^m-1}$ haciendo $\varepsilon_{\kappa_m(j)} = \delta_{j-m}$ para $j \in [\![m+1, 2^m-1]\!]$ y, para $j \in [\![1, m]\!]$, los valores $\varepsilon_{\kappa_m(j)}$ quedan determinados por la ec. (7).

Para decodificar un $\boldsymbol{\varepsilon} = (\varepsilon_j)_{j=1}^{2^m-1}$, se revisa si éste está en el código. El vector $H_m^{\perp}\boldsymbol{\varepsilon}$ se dice ser el *síndrome* de $\boldsymbol{\varepsilon}$. Si el síndrome fuese nulo no se hace ninguna corrección y se recupera $\delta \in \mathbb{F}_2^{2^m-1-m}$ mediante los bits de información: $\delta_j = \varepsilon_{\kappa_m(j+m)}$, para $j \in [\![1, 2^m-1-m]\!]$. En otro caso, deben existir $\boldsymbol{\varepsilon}'$ en el código de Hamming y un índice $i \in [\![1, 2^m-1]\!]$ tales que $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}' + \mathbf{e}_i$. Por tanto, ha de tenerse $H_m^{\perp}\boldsymbol{\varepsilon}' = \mathbf{0}$ y

$$H_m^{\perp}\boldsymbol{\varepsilon} = H_m^{\perp}(\boldsymbol{\varepsilon}' + \mathbf{e}_i) = H_m^{\perp}\mathbf{e}_i = \left(i\text{-ésima columna de } H_m^{\perp}\right)$$

y la $i$-ésima columna de $H_m^{\perp}$ no es otra que la representación en base 2 del índice $i$. Así pues, el síndrome indica cuál es el índice que ha de conmutarse para corregir el error.

**Observación 4.4** *Los códigos de Hamming poseen peso mínimo* 3.

En efecto, si la matriz revisora de paridad se escribe como las representaciones en base 2 de los números en $[\![1, 2^m-1]\!]$, entonces las tres primeras columnas, correspondientes a 1, 2 y 3 forman una submatriz $m \times 3$ con un bloque inicial de $(m-2)\times 3$ ceros y los dos últimos renglones son $\begin{matrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{matrix}$. De aquí se ve que la palabra $1110^{(2^m-4)}$ está en el código. Por tanto $w_m(H_m) \leq 3$. Por otro lado, como el código corrige un bit, por la observación 4.1, se tiene que $w_m(H_m) \geq 3$. □

**Observación 4.5** *Los códigos de Hamming son* perfectos *en el sentido de que cualquier palabra en el espacio que contiene a las palabras de código, es bien una palabra de código o bien dista en* 1 *de una palabra de código.*

**Definición 4.8** *Supongamos que $C \subset \mathbb{F}_2^n$ es un código lineal con palabras de código de longitud $n$, y que un canal binario simétrico, con probabilidad $p$ de alterar el valor de cada bit, transmite una palabra de código $\boldsymbol{\varepsilon}$. Si $\boldsymbol{\varepsilon}'$ es la palabra recibida tras la transmisión, el* error *es $\boldsymbol{\eta} = \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}'$. El error quedará* indetectado *siempre que $\boldsymbol{\eta} \in C$.*

Para cada $i \in [\![1, 2^m-1]\!]$, sea $c_i$ el número de palabras con peso de Hamming $i$ en el código $C$. Entonces, la probabilidad de que un error quede indetectado será

$$\text{Prob}\,(\text{error indetectado}) = P_x = \sum_{i=1}^{n} c_i p^i (1-p)^{n-i} = \sum_{i=d}^{n} c_i p^i (1-p)^{n-i}$$

donde $d$ es el peso mínimo de $C$ (si $i < d$ entonces $c_i = 0$).

**Definición 4.9** *El polinomio $P_C(X) = \sum_{i=0}^{n} c_i X^i$, donde $c_i$ es el número de palabras con peso $i$ en $C$, se dice ser el* enumerador de pesos *del código $C$.*

De acuerdo con lo anterior, se tiene

$$P_x = (1-p)^n \left[ C\left(\frac{p}{1-p}\right) - 1 \right].$$

Por ejemplo, para los primeros códigos de Hamming se tiene:

$m = 3$. La dimensión del código es $2^m - 1 - m = 8 - 1 - 3 = 4$, por tanto el código posee $2^4 = 16$ palabras, que clasificadas según sus pesos de Hamming producen las siguientes cuentas:

| $i$ | 0 | 3 | 4 | 7 |
|---|---|---|---|---|
| $c_i$ | 1 | 7 | 7 | 1 |

El enumerador de pesos es pues

$$P_C(X) = 1 + 7X^3 + 7X^4 + X^7.$$

$m = 4$. La dimensión del código es $2^m - 1 - m = 16 - 1 - 4 = 11$, por tanto el código posee $2^{11} = 2048$ palabras, que clasificadas según sus pesos de Hamming producen las siguientes cuentas:

| $i$ | 0 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_i$ | 1 | 35 | 105 | 168 | 280 | 435 | 435 | 280 | 168 | 105 | 35 | 1 |

El enumerador de pesos es pues

$$P_C(X) = 1+35X^3+105X^4+168X^5+280X^6+435X^7+435X^8+280X^9+168X^{10}+105X^{11}+35X^{10}+X^{15}.$$

$m = 5$. La dimensión del código es $2^m - 1 - m = 32 - 1 - 5 = 26$, por tanto el código posee $2^{26} = 67\,108\,864$ palabras, que clasificadas según sus pesos de Hamming producen las siguientes cuentas:

| $i$ | $c_i$ |
|---|---|
| 0 | 1 |
| 3 | 155 |
| 4 | 1085 |
| 5 | 5208 |
| 6 | 22568 |
| 7 | 82615 |
| 8 | 247845 |
| 9 | 628680 |
| 10 | 1383096 |
| 11 | 2648919 |
| 12 | 4414865 |
| 13 | 6440560 |
| 14 | 8280720 |
| 15 | 9398115 |

| $i$ | $c_i$ |
|---|---|
| 31 | 1 |
| 28 | 155 |
| 27 | 1085 |
| 26 | 5208 |
| 24 | 82615 |
| 25 | 22568 |
| 23 | 247845 |
| 22 | 628680 |
| 21 | 1383096 |
| 20 | 2648919 |
| 19 | 4414865 |
| 18 | 6440560 |
| 17 | 8280720 |
| 16 | 9398115 |

Si se conoce el polinomio enumerador de pesos en un código-$(n, k)$ $C$ se puede calcular procedimentalmente el polinomio enumerador de pesos de su dual $C^\perp$. Denotemos por $P_C(X)$ al enumerador de pesos de $C$, según la definición 4.9. Definamos al polinomio de dos variables

$$Q_C(X, Y) = Y^n \, P_C\left(\frac{X}{Y}\right) = \sum_{i=0}^{n} c_i X^i Y^{n-i}.$$

**Teorema 4.1 (MacWilliams)** *Para el código dual $C^\perp$ de $C$ se tiene*

$$P_{C^\perp}(X) = 2^{-k}(1 + X)^n \; P_C\left(\frac{1 - X}{1 + X}\right)$$

*o equivalentemente*

$$Q_{C^\perp}(X, Y) = 2^{-k} \; Q_C\left(Y - X, Y + X\right). \tag{8}$$

Observamos primero

$$2^{-k}\sum_{\mathbf{v}\in C}(-1)^{\langle \mathbf{u}|\mathbf{v}\rangle} = \begin{cases} 1 & \text{si } \mathbf{u}\in C^{\perp} \\ 0 & \text{si } \mathbf{u}\notin C^{\perp} \end{cases} \tag{9}$$

En efecto, por un lado $\operatorname{card} C = 2^k$. Por otro, si $\mathbf{u}\in C^{\perp}$ entonces $\langle \mathbf{u}|\mathbf{v}\rangle = 0$. En otro caso, existe $\mathbf{v}_0 \in C$ tal que $\langle \mathbf{u}|\mathbf{v}_0\rangle \neq 0$, es decir, $\langle \mathbf{u}|\mathbf{v}_0\rangle = 1$. Se puede ver que $\left(V\cap \mathbf{u}^{\perp}\right)$ y $\mathbf{v}_0 + \left(V\cap \mathbf{u}^{\perp}\right)$ forman una partición de $V$ y ambas tienen una misma cardinalidad. De aquí se sigue (9).

Observamos luego que, para cualquier código, se tiene

$$Q_C(X,Y) = \sum_{\mathbf{v}\in C} X^{|\mathbf{v}|}Y^{n-|\mathbf{v}|}.$$

Pues bien, por un lado se tiene

$$\begin{aligned} Q_{C^{\perp}}(X,Y) &= \sum_{\mathbf{u}\in C^{\perp}} X^{|\mathbf{u}|}Y^{n-|\mathbf{u}|} \\ &= \sum_{\mathbf{u}\in \mathbb{F}_2^n}\left[2^{-k}\sum_{\mathbf{v}\in C}(-1)^{\langle \mathbf{u}|\mathbf{v}\rangle}\right] X^{|\mathbf{u}|}Y^{n-|\mathbf{u}|} \\ &= 2^{-k}\sum_{\mathbf{v}\in C}\sum_{\mathbf{u}\in \mathbb{F}_2^n}(-1)^{\langle \mathbf{u}|\mathbf{v}\rangle} X^{|\mathbf{u}|}Y^{n-|\mathbf{u}|} \\ &= 2^{-k}\sum_{\mathbf{v}\in C}\sum_{i=0}^{n}\left[\sum_{w(\mathbf{u})=i}(-1)^{\langle \mathbf{u}|\mathbf{v}\rangle}\right] X^{i}Y^{n-i} \\ &= \sum_{\mathbf{v}\in C}\prod_{j=0}^{n-1}\left(Y+(-1)^{v_j}X\right). \end{aligned}$$

Por otro lado,

$$\begin{aligned} Q_C(Y-X,Y+X) &= \sum_{\mathbf{v}\in C}(Y-X)^{|\mathbf{v}|}(Y+X)^{n-|\mathbf{v}|} \\ &= \sum_{\mathbf{v}\in C}(Y-X)^{\sum_{j=0}^{n-1}v_j}(Y+X)^{n-\sum_{j=0}^{n-1}v_j} \\ &= \sum_{\mathbf{v}\in C}\prod_{j=0}^{n-1}\left(Y+(-1)^{v_j}X\right) \end{aligned}$$

Por tanto se cumple (8). □

# 5 Códigos lineales

## 5.1 Matrices generatrices, de paridad y sistemáticas

Sea $\mathbb{K}$ un campo cualquiera finito, de cardinalidad, digamos, $q$, el cual es, por consiguiente, una potencia de la característica de $\mathbb{K}$. Para cada $n \geq 0$ denotemos por $\mathbb{K}^n$ a su $n$-ésima potencia cartesiana dotada de su estructura usual de espacio vectorial sobre el campo $\mathbb{K}$.

**Definición 5.1** *Todo subespacio $C < \mathbb{K}^n$ de dimensión $k$ es un* código lineal-$(n,k)$ *sobre el alfabeto $\mathbb{K}$.*

Todo código lineal-$(n,k)$ tiene $k$ *símbolos de información* y $n-k$ *símbolos de revisión.* Está conformado por $q^k$ *palabras de código.*

Si $\{\mathbf{a}_j\}_{j=0}^{k-1} \subset \mathbb{K}^n$ es una base de $C$, la matriz $A = [\mathbf{a}_j]_{j=0}^{k-1} \in \mathbb{K}^{n\times k}$, cuyas columnas son los vectores en la base, se dice ser *generatriz* del código $C$. En tal caso, la transformación $A : \mathbb{K}^k \to \mathbb{K}^n$, $\mathbf{x}\mapsto \mathbf{y} = A\mathbf{x}$,

tiene como imagen a $C$ precisamente. Mediante la aplicación de una transformación lineal de permutación $P$ en el espacio $\mathbb{K}^n$, se tiene $A_s = PA$ donde $A_s$ es una matriz generatriz de la forma $A_s = \begin{bmatrix} I_k \\ A_r \end{bmatrix}$, con $A_r \in \mathbb{K}^{(n-k)\times k}$, llamada *sistemática*.

**Definición 5.2** *Dos códigos $C$ y $D$ se dicen* equivalentes *si una generatriz de uno se obtiene mediante la aplicación de una matriz de permutación a la generatriz del otro.*

Así todo código lineal es equivalente a otro con una generatriz sistemática.

**Definición 5.3** *Una matriz $B \in \mathbb{K}^{(n-k)\times n}$ es* revisora de paridad *para un código lineal-$(n,k)$ $C < \mathbb{K}^n$ si se cumple:*

$$\forall \mathbf{y} \in \mathbb{K}^n : \ [\mathbf{y} \in C \iff B\mathbf{y} = \mathbf{0}] .$$

Se tiene, de manera natural:

$$\begin{bmatrix} I_k \\ A_r \end{bmatrix} \text{ generatriz } \iff [-A_r \ \ I_{n-k}] \text{ revisora de paridad.}$$

Por tanto se tiene que toda matriz revisora de paridad posee rango $n-k$.

**Definición 5.4** *Dado un código lineal-$(n,k)$ $C < \mathbb{K}^n$, su* código dual *es $C^\perp = \{\mathbf{y} \in \mathbb{K}^n | \ \langle \mathbf{y}|\mathbf{x}\rangle = 0 \ \forall \mathbf{x} \in C\}$.*

Naturalmente, la matriz revisora de paridad de un código es la generatriz del dual y, viceversa, la generatriz es la revisora de paridad del dual.

**Definición 5.5** *Si $B \in \mathbb{K}^{(n-k)\times n}$ es la matriz revisora de paridad de un código lineal-$(n,k)$ $C < \mathbb{K}^n$ entonces la transformación lineal $B : \mathbb{K}^n \to \mathbb{K}^{n-k}$, $\mathbf{y} \mapsto B\mathbf{y}$, se dice ser* de síndrome. *El valor $B\mathbf{y}$ es el* síndrome *de la palabra $\mathbf{y}$.*

Así tenemos que las palabras en el código son exactamente aquellas con síndrome nulo. En otras palabras, el núcleo de la transformación de síndrome es el código mismo.

## 5.2 Pesos mínimos

**Definición 5.6** *El* peso de Hamming *de una palabra $\mathbf{y} \in \mathbb{K}^n$ es $w(\mathbf{y}) = \text{card}\{j \in [\![0, n-1]\!] | \ y_j \neq 0\}$. Para un código lineal-$(n,k)$ $C < \mathbb{K}^n$, su* peso mínimo *es $w_m(C) = \min\{w(\mathbf{y}) | \ \mathbf{y} \in C - \{\mathbf{0}\}\}$.*

**Observación 5.1 (Cota de Singleton)** *El peso mínimo $w_m(C)$ de un código lineal-$(n,k)$ $C$ queda acotado como sigue:*

$$w_m(C) \leq n - k + 1. \tag{10}$$

En efecto, por un lado como los símbolos de información pueden asumir cualesquiera valores, transformando a una generatriz sistemática se ve que la palabra $\mathbf{y} = 1 \ 0^{k-1} \ y_{k+1} \cdots y_{n-1}$ es una palabra en el código de peso a lo sumo $n-(k-1)$, por tanto $w_m(C) \leq n-k+1$. □

Como es convencional, si $d = w_m(C)$, entonces se dice que $C$ es un *código lineal-$(n,k,d)$*.

Un código lineal-$(n,k,d)$ $C$ que alcance la cota de Singleton, es decir, en el que vale la igualdad en la relación (10), se dice ser *separable con la distancia máxima* (en inglés, *Maximum Distance Separable* (MDS)). Primeramente:

**Observación 5.2** *Un código lineal-$(n,k)$ $C$ es MDS cuando y sólo cuando las palabras en $C$ no-nulas de peso mínimo tienen peso $n-(k-1)$.*

Otra caracterización es la siguiente:

**Observación 5.3** *Un código lineal-$(n,k)$ $C$ es MDS cuando y sólo cuando cualesquiera $n-k$ columnas de una matriz revisora de paridad de $C$ son linealmente independientes.*

Y considerando códigos duales, una caracterización más es:

**Observación 5.4** *Sea $C$ un código lineal-$(n,k)$. Entonces las siguientes relaciones son equivalentes a pares:*

1. *$C$ es MDS.*
2. *El código dual $C^\perp$ es MDS.*
3. *Cualesquiera $k$ columnas de una matriz generatriz de $C$ son linealmente independientes.*
4. *Si $\begin{bmatrix} I_k \\ A_r \end{bmatrix}$, con $A_r \in \mathbb{K}^{(n-k)\times k}$, es una matriz sistemática de $C$ entonces cualquier submatriz cuadrada de $A_r$ es no-singular.*

Por otro lado:

**Observación 5.5** *Sea $B \in \mathbb{K}^{(n-k)\times n}$ una matriz revisora de paridad del código lineal-$(n,k,d)$ $C$. Entonces cualesquiera $d-1$ columnas de $B$ son linealmente independientes.*

En efecto supongamos que hubiese $m$ columnas de $B$ linealmente dependientes, con $m \leq d-1$.

Sea $M \subset [\![0, n-1]\!]$ el conjunto de índices correspondiente a esas columnas. Entonces sin pérdida de generalidad podemos suponer que la suma de esas columnas es el vector cero, $\sum_{j\in M} \mathbf{b}_j = \mathbf{0} \in \mathbb{K}^{n-k}$.

Sea $\mathbf{x} = \sum_{j\in M} \mathbf{e}_j \in \mathbb{K}^n$. Entonces $\mathbf{x}$ tiene peso $m$ y su síndrome es $B\mathbf{x} = \sum_{j\in M} \mathbf{b}_j$, o sea es el vector cero. Por tanto $\mathbf{x}$ ha de estar en el código $C$, pero esto no es posible ya que el peso mínimo de $C$ es $d$ y éste es mayor que $m$. □

Una reformulación de la observación 5.5 es la siguiente:

**Observación 5.6** *Sea $B \in \mathbb{K}^{(n-k)\times n}$ una matriz revisora de paridad del código lineal-$(n,k,d)$ $C$. Entonces $d$ coincide con el entero mínimo $m$ para el cual hay $m$ columnas linealmente dependientes en $B$.*

Consideremos ahora $\mathbb{K} = \mathbb{F}_2$. Sea $b(n,r) = \sum_{j=0}^{r} \binom{n}{j}$ la cardinalidad de una bola de radio $r$ centrada en un punto del hipercubo $\mathbb{F}_2^n$.

**Observación 5.7 (Cota de Hamming)** *Sea $C$ un código lineal-$(n,k,d)$. Entonces*

$$\log_2 b\left(n, \left\lfloor \frac{d-1}{2} \right\rfloor\right) \leq n-k. \tag{11}$$

En efecto, sea $r = \left\lfloor \frac{d-1}{2} \right\rfloor$. Para cualesquiera dos palabras de código $\mathbf{x}_0, \mathbf{x}_1 \in C$, al ser éste de peso mínimo $d$, se tiene $[\mathbf{x}_0 \neq \mathbf{x}_1 \;\Rightarrow\; B(\mathbf{x}_0, r) \cap B(\mathbf{x}_1, r) = \emptyset]$, es decir, cualesquiera dos bolas de radio $r$ con centros en palabras de código distintas han de ser ajenas. Por tanto,

$$2^n \geq \operatorname{card}\left(\bigcup_{\mathbf{x}\in C} B(\mathbf{x}, r)\right) = 2^k \cdot b\left(n, \left\lfloor \frac{d-1}{2} \right\rfloor\right),$$

de donde se sigue (11). □

Un código lineal-$(n,k,d)$ $C$ que alcance la cota de Hamming, es decir, en el que vale la igualdad en la relación (11), se dice ser *perfecto*. En un tal código, se tiene que $\left\{B\left(\mathbf{x}, \left\lfloor \frac{d-1}{2} \right\rfloor\right)\right\}_{\mathbf{x}\in C}$ es una partición del hipercubo $\mathbb{F}_2^n$. Los códigos de Hamming son perfectos. Naturalmente, la noción de perfección introducida aquí generaliza a la de la observación 4.5.

## 5.3 Arreglos estándares

Sea $\mathbb{K}$ un campo finito y sea $C$ un código lineal-$(n,k,d)$. Al ser $C$ un subespacio, el cociente $\mathbb{K}^n/C = \{\mathbf{y}+C \,|\, \mathbf{y} \in \mathbb{K}^n\}$ es a su vez un espacio vectorial sobre $\mathbb{K}$. Naturalmente, dos palabras cualesquiera $\mathbf{y}, \mathbf{z} \in \mathbb{K}^n$ en una misma clase del cociente $\mathbb{K}^n/C$, es decir, tales que $\mathbf{z}-\mathbf{y} \in C$, han de poseer el mismo síndrome: $B\mathbf{z} = B\mathbf{y}$ donde $B \in \mathbb{K}^{(n-k)\times n}$ es la matriz revisora de paridad de $C$.

También, si al transmitir una palabra en el código, digamos $\mathbf{y} \in C$, se recibiera la palabra $\mathbf{z} \in \mathbb{K}^n$ entonces para el error $\mathbf{e} = \mathbf{z}-\mathbf{y}$ se habría de tener $B\mathbf{e} = B\mathbf{z}$. Así pues, el síndrome del error ha de coincidir con el síndrome de la palabra recibida, lo que, por lo anterior, equivale a que la palabra recibida $\mathbf{z}$ y el error cometido $\mathbf{e}$ necesariamente han de estar en una misma clase lateral de $\mathbb{K}^n/C$.

**Definición 5.7** *Para cada clase $\mathbf{z}+C \in \mathbb{K}^n/C$, un* representante principal *de ella es un vector $\mathbf{e} \in \mathbf{z}+C$ de peso de Hamming mínimo.*

Por el Teorema Fundamental de Homomorfismos se tiene que $B : \mathbb{K}^n/C \to \mathrm{Img}(B)$ es un isomorfismo. Así, para cada posible valor de síndrome $\mathbf{s} \in \mathrm{Img}(B) < \mathbb{K}^{n-k}$ existe una única clase lateral $\mathbf{z}_s + C \in \mathbb{K}^n/C$ tal que $B(\mathbf{z}_s + C) = \mathbf{s}$. Sea $\mathbf{e}_s$ un representante principal de la clase $\mathbf{z}_s + C$. Resulta entonces un

**Procedimiento de decodificación.** Supóngase que al transmitir una palabra $\mathbf{y} \in C$ se recibe la palabra $\mathbf{z} \in \mathbb{K}^n$ cometiéndose el error $\mathbf{e} = \mathbf{z}-\mathbf{y}$. Entonces se calcula el síndrome $\mathbf{s} = B\mathbf{z}$ y, considerando el representante principal $\mathbf{e}_s$, se recupera la palabra transmitida tomando $\mathbf{z}-\mathbf{e}_s$.

Este procedimiento es, claramente, correcto toda vez que $\mathbf{e} = \mathbf{e}_s$. Una manera de sistematizarlo es la siguiente:

**Definición 5.8** *El* arreglo estándar *del espacio $\mathbb{K}^n$, mediante el código $C$, es la matriz*

$$K_n = [\mathbf{y}_{ij}]_{i\in[\![0,q^{n-k}-1]\!]}^{j\in[\![0,q^k-1]\!]}$$

*tal que*

- *cada renglón es una clase, i.e. $\forall i \in [\![0, q^{n-k}-1]\!]\ \exists \mathbf{y}_i \in \mathbb{K}^n$: $\{\mathbf{y}_{ij}\}_{j\in[\![0,q^k-1]\!]} = \mathbf{y}_i + C$,*
- *el primer renglón consiste de las palabras de código, i.e. $\mathbf{y}_0 = \mathbf{0}$,*
- *la primera columna consta de representantes principales, i.e. $\mathbf{y}_{i0}$ es de peso mínimo en $\mathbf{y}_i + C$, y*
- *en cada entrada aparece la suma del representante principal de la clase con la correspondiente palabra de código, i.e. $\forall i \in [\![0, q^{n-k}-1]\!], j \in [\![0, q^k-1]\!]$: $\mathbf{y}_{ij} = \mathbf{y}_{i0} + \mathbf{y}_{0j}$.*

Utilizando el arreglo estándar se tiene el siguiente:

**Procedimiento de decodificación.** Supóngase que al transmitir una palabra $\mathbf{y} \in C$ se recibe la palabra $\mathbf{z} \in \mathbb{K}^n$ cometiéndose el error $\mathbf{e} = \mathbf{z}-\mathbf{y}$. Entonces se localiza en $K_n$ los índices $i, j$ tales que $\mathbf{z} = \mathbf{y}_{ij}$, y se toma a $\mathbf{y}_{0j} \in C$ como la palabra original $\mathbf{y}$.

Este procedimiento es, claramente, correcto toda vez que el error $\mathbf{e}$ coincida con el representante principal $\mathbf{y}_{i0}$ de la clase $\mathbf{y}_i + C$ en la que apareció $\mathbf{z}$.

## 5.4 Segundas listas

### 5.4.1 Ejercicios

1. Verifique que las tablas siguientes definen una estructura de campo en el conjunto de 4 elementos $F = \{0, 1, p, q\}$:

| $+$ | $0$ | $1$ | $p$ | $q$ |
|---|---|---|---|---|
| $0$ | $0$ | $1$ | $p$ | $q$ |
| $1$ | $1$ | $0$ | $q$ | $p$ |
| $p$ | $p$ | $q$ | $0$ | $1$ |
| $q$ | $q$ | $p$ | $1$ | $0$ |

| $\cdot$ | $0$ | $1$ | $p$ | $q$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $p$ | $q$ |
| $p$ | $0$ | $p$ | $q$ | $1$ |
| $q$ | $0$ | $q$ | $1$ | $p$ |

Indique cómo se puede extender esas tablas de manera que determinen un campo de 8 elementos.

2. El polinomio $p(X) = 1+X^2+X^5$ es irreducible en $\mathbb{F}_2[X]$. Describa la tabla de adición y de multiplicación del cociente $\mathbb{F}_2[X]/(p(X))$.

3. Demuestre que el conjunto $\mathbb{K}^{m\times n}$, de matrices de orden $m \times n$ y entradas en un campo $\mathbb{K}$, es un espacio vectorial sobre $\mathbb{K}$. Determine la dimensión de $\mathbb{K}^{m\times n}$.

4. Sea $C$ un código-$[n, k]$ binario. Suponga que la distancia mínima de $C$ es bien $2t$ o bien $2t + 1$ y que se está usando un canal simétrico con probabilidad de error $p$. Muestre que la probabilidad de error en el código satisface $P_{err}(C) \leq \sum_{j=t+1}^{n} \binom{n}{j} p^j (1-p)^{n-j}$.

5. Suponga un canal simétrico con probabilidad de error $p = 10^{-1}$. ¿Cuál debería ser el tamaño de $n$ para que en el código de $n$ repeticiones de cada bit se tenga $P_{err}(C) \leq 10^{-4}$?

6. Si $\sigma = s_0 \cdots s_{n-1} \in \{0, 1\}^*$ es una palabra binaria su *reverso* es la palabra $\text{rev}(\sigma) = s_{n-1} \cdots s_0 \in \{0, 1\}^*$. El conjunto de *n-palíndromas* es $P_n = \{\sigma \in \{0, 1\}^n | \, \sigma = \text{rev}(\sigma)\}$ que consta de las palabras de longitud $n$ que se leen iguales en cualquier sentido. Decida si $P_n$ es un código lineal, y si lo fuera determine un conjunto de ecuaciones para determinar cuándo una palabra cae en el código. Calcule la distancia mínima de este código y diga cuántos errores de bits puede detectar.

7. Describa un procedimiento para detectar errores triples cuando se utiliza un código recangular $5 \times 2$.

8. Sea $\mathbb{F}_p^n$ el espacio de dimensión $n$ sobre el campo $\mathbb{F}_p$, donde $p$ es un número primo. Para cada entero $k \leq n$ cuente cuántos subespacios de dimensión $k$ hay en $\mathbb{F}_p^n$.

9. Pruebe que la decodificación de un código de Hamming es siempre incorrecta si hay dos bits erróneos en una misma palabra de código.

10. Considere el código lineal $C \subset \mathbb{F}_2^n$ determinado por el sistema de ecuaciones

$$\begin{aligned} X_4 &= X_1 + X_2 + X_3 \\ X_5 &= X_0 + X_1 + X_2 \\ X_6 &= X_0 + X_1 + X_4 \\ X_7 &= X_0 + X_2 + X_3 \end{aligned}$$

Calcule una matriz de paridad y verifique que la distancia mínima es 3.

11. En el espacio $\mathbb{F}_2^n$ sea $C_0$ el código consistente de los vectores con peso de Hamming par. Encuentre una matriz generatriz y una revisora de paridad y calcule la distancia mínima.

12. ¿Cómo son los códigos en $\mathbb{F}_2^n$ cuyas matrices generatrices son invertibles?

13. Muestre que los triángulos equiláteros en un espacio vectorial $\mathbb{F}_2^n$ tienen aristas pares. En otras palabras, muestre que si $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{F}_2^n$ son tales que la distancia entre cualesquiera dos de ellos es $d \in \mathbb{N}$, entonces $d$ es un número par. Muestre que en tal caso hay un único punto $\mathbf{c} \in \mathbb{F}_2^n$ cuya distancia a cada $\mathbf{x}_i$ es $d/2$. Es decir, todo triángulo equilátero posee un centro.

14. Sea $C$ un código lineal binario. Muestre que bien todas las palabras en $C$ comienzan con 0 o bien exactamente una mitad del código consta de palabras que comienzan con 0.

15. En el espacio $\mathbb{F}_2^n$ sea $C_1 = \{0^n, 1^n\}$ el código con sólo dos palabras: las constantes 0 y 1. Encuentre una matriz generatriz y una revisora de paridad y calcule la distancia mínima.

16. Sea $C$ un código lineal en el espacio $\mathbb{F}_2^n$. Para cada $\mathbf{x} \in \mathbb{F}_2^n$, denotemos por $\mathbf{x} + C$ a la clase lateral en $\mathbb{F}_2^n/C$ que contiene a $\mathbf{x}$.

Muestre que si $\mathbf{y} \in \mathbf{x} + C$ entonces $\mathbf{y} + C = \mathbf{x} + C$, y si $\mathbf{y} \notin \mathbf{x} + C$ entonces $(\mathbf{y} + C) \cap (\mathbf{x} + C) = \emptyset$.

17. Sea $C$ un código lineal en el espacio $\mathbb{F}_2^n$. Muestre que $C \cup (\mathbf{x} + C)$ también es un código lineal en el espacio $\mathbb{F}_2^n$.

18. Pruebe que los códigos-$(2^n - 1, 2^n - n - 1, 3)$ de Hamming son perfectos.

19. Un código lineal-$(n, k)$ $C$ se dice ser *autodual* si $C = C^\perp$. Muestre que un código lineal-$(n, k)$ $C$ en el espacio $\mathbb{F}_2^n$ es autodual si y sólo si cualesquiera dos renglones en una matriz generatriz de $C$ son ortogonales (su producto interno es cero) y $k = \frac{n}{2}$.

20. Sea $H \in \mathbb{F}_2^{(n-k)\times n}$ la matriz revisora de paridad de un código-$[n,k]$ $C$ cuyo peso mínimo es un entero impar. Construya un nuevo código $D$ cuya matriz revisora de paridad es

$$\overline{H} = \begin{bmatrix} H & 1^{(n-k)} \\ \left(0^{(n)}\right)^T & 1 \end{bmatrix}.$$

Muestre que $D$ es un código-$[n+1,k]$ cuyo peso mínimo es $d+1$.

### 5.4.2 Programas

1. Escriba un programa que lea de un archivo ASCII el valor de un número primo $p$, de un entero $n$, y una matriz $A$ de orden $n \times n$ con entradas en $\mathbb{F}_p$, decida si $A$ es invertible en $\mathbb{F}_p$ y en caso de que lo sea calcule $A^{-1}$.

2. Escriba un programa que lea de un archivo ASCII el valor de un número primo $p$, de dos enteros $m, n$, una matriz $A$ de orden $m \times n$ con entradas en $\mathbb{F}_p$, y un vector $b$ de dimensión $m$ con entradas en $\mathbb{F}_p$, decida si el sistema de ecuaciones $Ax = b$ posee soluciones $x \in \mathbb{F}_p^n$ y en caso de que las haya, las describa.

3. Escriba un programa que lea de un archivo ASCII el valor de un número primo $p$, de dos enteros $k, n$, una matriz $A$ de orden $n \times k$ con entradas en $\mathbb{F}_p$, generatriz de un código-$[n,k]$, y calcule una matriz de paridad $A^{\perp} \in \mathbb{F}_p^{(n-k)\times n}$.

4. Escriba un programa que reciba un valor de probabilidad $p < \frac{1}{2}$, correspondiente a que un cierto canal binario simétrico altere el valor de un bit, y un entero impar $n$, longitud de un código por mayoría de votos, y contenga los procedimientos siguientes:

   A. Dado un bit $\varepsilon \in \{0,1\}$ genere la palabra $\delta = \delta_0 \cdots \delta_{n-1} \in \{0,1\}^n$ de la siguiente manera: Para cada $i \leq n-1$ genere un valor aleatorio real $r$ entre 0 y 1, si $r \geq p$ entonces debe hacer $\delta_i = \varepsilon$ y, en otro caso, $\delta_i = 1 - \varepsilon$ .

   B. Dada una sucesión de bits $\varepsilon = \varepsilon_0 \cdots \varepsilon_{m-1} \in \{0,1\}^m$ codifique cada uno como en el punto A. y obtenga un código $\delta \in \{0,1\}^{mn}$.

   C. Dada una palabra $\delta \in \{0,1\}^{mn}$ recupere $\varepsilon \in \{0,1\}^m$ tomando en cada bloque de $n$ bits contiguos aquel que aparezca más veces.

   D. Dado $\varepsilon \in \{0,1\}^m$, genere el código correspondiente $\delta \in \{0,1\}^{mn}$, según B., a éste aplique C. para obtener $\varepsilon' \in \{0,1\}^m$, y cuente las discrepancias entre $\varepsilon$ y $\varepsilon'$.

   E. Genere muchas, muchas pero muchas palabras aleatorias $\varepsilon \in \{0,1\}^m$ y tome el promedio de las discrepancias. Este valor depende de $p$ y de $n$.

   F. Fije $p$ y grafique estadísticas de desempeño variando $n$.

   G. Explique los resultados del programa y las observaciones de usted.

5. Escriba un programa que reciba dos enteros $m, n$, y contenga los procedimientos siguientes para manejar códigos rectangulares $m \times n$:

   A. Dada una palabra $\varepsilon \in \{0,1\}^{(m-1)(n-1)}$ calcule su código $\delta \in \{0,1\}^{mn}$ de acuerdo con el código rectangular $m \times n$.

   B. Encuentre un par de palabras en el código rectangular $m \times n$ cuya distancia de Hamming sea 4.

   C. Dada una palabra $\delta \in \{0,1\}^{mn}$, localice la palabra $\delta_0$ en el código rectangular $m \times n$ más cercana a $\delta$ y dé esa distancia.

   D. Realice el procedimiento de decodificación. En el caso de corregir errores, decida si es posible identificar la posición de los bits erróneos.

6. Escriba un programa que reciba la matriz generatriz y una correspondiente de paridad de un código-$[n, k]$ lineal, y contenga los procedimientos siguientes para manejar códigos lineales:

A. Dada una palabra $\varepsilon \in \{0,1\}^k$ calcule su código $\delta \in \{0,1\}^n$.

B. Encuentre un par de palabras en el código lineal cuya distancia de Hamming sea la distancia mínima del código.

C. Dada una palabra $\delta \in \{0,1\}^n$, localice la palabra $\delta_0$ en el código lineal más cercana a $\delta$ y dé esa distancia.

D. Realice el procedimiento de decodificación. En el caso de corregir errores, decida si es posible identificar la posición de los bits erróneos.

7. Escriba un programa que reciba $n$ y genere todos los códigos lineales binarios separables con la distancia máxima, escribiendo cada uno en un archivo ASCII distinto.
Conjeture cómo caracterizarlos y demuestre que su conjetura es válida.

8. Escriba un programa que reciba una matriz $G \in \mathbb{F}_2^{n\times k}$, generatriz de un código $C$ y realice las funciones siguientes:

A. Enliste todas las palabras del código $C$.

B. Transforme $G$ a una matriz sistemática equivalente.

C. Calcule una correspondiente matriz revisora de paridad.

D. Construya el arreglo estándar de $C$.

9. Escriba un programa que reciba una matriz $R \in \mathbb{F}_2^{n\times k}$, revisora de paridad de un código $C$ y realice las funciones siguientes:

A. Enliste todas las palabras del código $C$.

B. Calcule una correspondiente matriz generatriz $G$.

C. Transforme $G$ a una matriz sistemática equivalente.

D. Construya el arreglo estándar de $C$.

10. Escriba un programa que reciba $n, k \in \mathbb{N}$ y genere todas las matrices generatrices de códigos-$(n, k)$ que son MDS. Pare esto utilice la caracterización formulada en la observación 5.4.

# 6 Modificaciones de códigos

Sea $\mathbb{K}$ un campo finito y sea $C < \mathbb{K}^n$ un código lineal-$(n, k)$. Sea $C_e < \mathbb{K}^{n+1}$ definido como sigue:

$$\forall(\mathbf{x}, x_n) \in \mathbb{K}^n \times \mathbb{K} : \left[ (\mathbf{x}, x_n) \in C_e \iff \mathbf{x} \in C \ \&\ x_n = \sum_{j=0}^{n-1} x_j \right].$$

Naturalmente, $C_e$ es un código lineal-$(n + 1, k)$. Si $B \in \mathbb{K}^{(n-k)\times n}$ es una matriz revisora de paridad de $C$ entonces una matriz revisora de paridad de $C_e$ es

$$B_e = \begin{bmatrix} B & 0_{n-k,1} \\ 1_{1n} & 1 \end{bmatrix}$$

donde $0_{n-k,1} \in \mathbb{K}^{(n-k)\times 1}$ es el vector columna consistente de $n - k$ ceros y $1_{1n} \in \mathbb{K}^{1\times n}$ es el vector renglón consistente de $n$ unos. $C_e$ se dice ser el código *extendido* de $C$.

**Observación 6.1** *Si $\mathbb{K} = \mathbb{F}_2$ y $C < \mathbb{K}^n$ es un código lineal-$(n,k)$ con peso mínimo $w_C$ impar, entonces el peso mínimo de $C_e$ es $w_{C_e} = w_C + 1$.*

En efecto, si $\mathbf{e}$ es de peso mínimo en $C$ entonces el correspondiente vector $(\mathbf{e}, e_n)$ en $C_e$ ha de ser tal que $e_n = 1$. Por tanto $w(\mathbf{e}, e_n) = w(\mathbf{e}) + 1$. □

**Observación 6.2** *Para los códigos de Hamming $H_m$, de acuerdo con la observación 4.4, sus extendidos $H_{me}$ son de peso mínimo 4, y una matriz revisora de paridad de ellos es de la forma*

$$H_{me}^{\perp} = \begin{bmatrix} H_m^{\perp} & 0_{m1} \\ 1_{1,2^m-1} & 1 \end{bmatrix}$$

Sea $C_r < \mathbb{K}^{n-1}$ definido como sigue:

$$\forall \mathbf{x} \in \mathbb{K}^{n-1} : \ [\mathbf{x} \in C_r \iff \exists x_n \in \mathbb{K} : \ (\mathbf{x}, x_n) \in C] .$$

Si se escribe a una matriz revisora de paridad de $C$ como

$$B = \begin{bmatrix} B_{00} & \mathbf{b}_{01} \\ \mathbf{b}_{10}^T & b_{11} \end{bmatrix}$$

entonces una matriz revisora de paridad de $C_r$ ha de ser $B_r = \left(\frac{1}{b_{11}} \mathbf{b}_{10}^T \mathbf{b}_{01} - B_{00}\right)$. El código $C_r$ se dice ser *recortado* de $C$.

**Observación 6.3** *Un código es el recortado de su extendido y es también equivalente al extendido de su recortado.*

Ahora, sea $C_a = C \oplus (1_n + C)$ el espacio generado por $C$ y la clase lateral $1_n + C$, donde $1_n$ es el vector en $\mathbb{K}^n$ constante 1:

$$\forall \mathbf{x} \in \mathbb{K}^n : \ [\mathbf{x} \in C_a \iff \exists \mathbf{y} \in C, x \in \mathbb{K} : \ \mathbf{x} = \mathbf{y} + x\, 1_n] .$$

El código $C_a$ se dice ser *aumentado* de $C$.

**Proposición 6.1** *Sea $\mathbb{K} = \mathbb{F}_2$. En todo código lineal ocurre que bien todas sus palabras poseen peso par, o bien el número de las palabras con peso par coincide con el número de las palabras con peso impar.*

En efecto, supóngase que hubiese una palabra de código $\mathbf{y} \in C$ de peso impar. Por un lado, como $C$ es un grupo abeliano con la suma, se tiene que la traslación $\mathbf{x} \mapsto \mathbf{y} + \mathbf{x}$ es una biyección. Por otro lado, para cualquier palabra de código $\mathbf{x}$ vale que el peso de $\mathbf{x}$ es par si y sólo si el peso de $\mathbf{y} + \mathbf{x}$ es impar. Por tanto, la mitad de los elementos de $C$ posee pesos pares. □

Supondremos en lo sucesivo que $\mathbb{K} = \mathbb{F}_2$.

Sea $C < \mathbb{F}_2^n$ un código lineal. Sea $C_x = \{\mathbf{y} \in C | w(\mathbf{y}) \equiv 0 \bmod 2\}$ el conjunto de palabras de código con peso par. Entonces $C_x$ es un código lineal, llamado *expurgado* de $C$. Por la proposición anterior, resulta que bien $C_x$ coincide con $C$ o bien posee la mitad de elementos de $C$.

**Definición 6.1** *Se dice que un código lineal $C < \mathbb{F}_2^n$* corrige $t$ errores y detecta $s$ errores simultáneamente *si*

$$\forall \mathbf{y} \in C \, \forall \mathbf{w} \in \mathbb{F}_2^n : \ [d_n(\mathbf{y}, \mathbf{w}) \leq s \implies \forall \mathbf{z} \in C - \{\mathbf{y}\} : \ d_n(\mathbf{w}, \mathbf{z}) > t] . \tag{12}$$

Con un tal código se puede decodificar según el siguiente:

**Procedimiento de decodificación.** Supóngase que al transmitir una palabra $\mathbf{y} \in C$ se recibe la palabra $\mathbf{w} \in \mathbb{F}_2^n$. Calcúlese la palabra de código $\mathbf{z} \in C$ más cercana a $\mathbf{w}$. Si $d_n(\mathbf{z}, \mathbf{w}) \leq t$ entonces dése a $\mathbf{z}$ como la palabra $\mathbf{y}$. En otro caso, anúnciese que al menos $s$ símbolos han cambiado.

**Proposición 6.2** *Un código lineal $C < \mathbb{F}_2^n$ corrige $t$ errores y detecta $s$ errores simultáneamente si y sólo si su distancia mínima $w_m(C)$ satisface*

$$w_m(C) \geq t + s + 1. \tag{13}$$

Supongamos primero que vale la desigualdad (13). Sean $\mathbf{y} \in C$ y $\mathbf{w} \in \mathbb{F}_2^n$ tales que $d_n(\mathbf{y}, \mathbf{w}) \leq s$. Para cualquier otra palabra $\mathbf{z} \in C - \{\mathbf{y}\}$ se tiene $d_n(\mathbf{z}, \mathbf{y}) \geq w_m(C) \geq t + s + 1$. Por la desigualdad del triángulo se sigue: $d_n(\mathbf{z}, \mathbf{w}) + d_n(\mathbf{w}, \mathbf{y}) \geq t + s + 1$; y por tanto

$$d_n(\mathbf{z}, \mathbf{w}) \geq t + s + 1 - d_n(\mathbf{w}, \mathbf{y}) \geq t + 1,$$

con lo que queda demostrada la relación (12).

Recíprocamente, supongamos que la desigualdad (13) no se cumpliera. Entonces $w_m(C) < t + s + 1$ y habría $\mathbf{y}, \mathbf{z} \in C$ tales que $d_n(\mathbf{y}, \mathbf{z}) = w_m(C) \leq t + s$. Elijamos $s$ posiciones donde las entradas de $\mathbf{y}$ y $\mathbf{z}$ difieran y sea $\mathbf{w} \in \mathbb{F}_2^n$ tal que coincida con $\mathbf{y}$ salvo en las $s$ posiciones seleccionadas, donde ha de tomar los valores de $\mathbf{z}$. Entonces $d_n(\mathbf{y}, \mathbf{w}) = s$ y $d_n(\mathbf{z}, \mathbf{y}) = d_n(\mathbf{z}, \mathbf{w}) + d_n(\mathbf{y}, \mathbf{w})$. Por tanto,

$$d_n(\mathbf{z}, \mathbf{w}) = d_n(\mathbf{z}, \mathbf{y}) - d_n(\mathbf{y}, \mathbf{w}) \leq (t + s) - s = t$$

lo cual evidencia que la relación (12) tampoco puede cumplirse. □

De la observación 6.2 se tiene que el extendido del código de Hamming posee un peso mínimo $w_m(H_{me}) = 4 \geq 2 + 1 + 1$, por tanto $H_{me}$ puede corregir un error y detectar dos errores simultáneamente. Se puede pues decodificar como sigue:

**Procedimiento de decodificación.** Supóngase que al transmitir una palabra $(\mathbf{y}, y_{2^m}) \in H_{me}$ se recibe la palabra $(\mathbf{w}, w_{2^m}) \in \mathbb{F}_2^{2^m}$. Calcúlese el síndrome $\mathbf{s} = H_m^{\perp}\mathbf{w}$. Si $\mathbf{s} = \mathbf{0}$ entonces acéptese $\mathbf{w}$ como $\mathbf{y}$ y acaso corríjase $w_{2^m}$ si fuera necesario; en otro caso, si $w_{2^m} = 0$ declárese que hubo al menos dos errores y si $w_{2^m} = 1$ entonces cámbiese el valor $w_i$ donde $i$ está dado en binario por el síndrome $\mathbf{s}$.

# 7 Códigos de Reed-Muller

## 7.1 Funciones booleanas

**Definición 7.1 (Funciones booleanas)** *Para cada $n \in \mathbb{N}$ sea $\mathbb{F}_2^{\mathbb{F}_2^n} = \{f : \mathbb{F}_2^n \to \mathbb{F}_2 | \, f \text{ es función}\}$ la colección de* funciones booleanas.

Naturalmente, $\mathbb{F}_2^{\mathbb{F}_2^n}$ posee una estructura de álgebra booleana con $\mathbf{0}$ como elemento mínimo y $\mathbf{1}$ como elemento máximo. También, visto como un espacio vectorial sobre su campo primo $\mathbb{F}_2$, la colección $\mathbb{F}_2^{\mathbb{F}_2^n}$ posee una estructura de espacio vectorial de dimensión $2^n$. Una base del espacio está dada por las funciones $(\delta_{\boldsymbol{\varepsilon}_0})_{\boldsymbol{\varepsilon}_0 \in \mathbb{F}_2^n}$ donde $\delta_{\boldsymbol{\varepsilon}_0} : \boldsymbol{\varepsilon}_1 \mapsto \delta_{\boldsymbol{\varepsilon}_0 \boldsymbol{\varepsilon}_1}$ y esta última es la delta de Kroenecker.

**Observación 7.1** *Las siguientas son identificaciones naturales entre respectivos conjuntos:*

- *La correspondencia $\iota_n : i \mapsto (i)_2$, que a cada entero le asocia su representación en base 2 de longitud $n$, identifica a $[\![0, 2^n - 1]\!]$ con $\mathbb{F}_2^n$.*
- *La correspondencia $I_n : f \mapsto [f(\iota_n(i))]_{i \in [\![0, 2^n - 1]\!]}$, que a cada función booleana le asocia la cadena de sus valores de acuerdo con el orden de $[\![0, 2^n - 1]\!]$, identifica a $\mathbb{F}_2^{\mathbb{F}_2^n}$ con $\mathbb{F}_2^{2^n}$. Mediante tal identificación se dice que toda función booleana es una* palabra-$2^n$.

Las funciones *proyecciones* $\pi_j : \boldsymbol{\varepsilon} \mapsto \varepsilon_j$, $j \in [\![0, n-1]\!]$, son funciones booleanas, y al formar un conjunto de $n$ funciones linealmente independientes, ellas mismas generan un espacio vectorial $\mathrm{Lin}(\mathbb{F}_2^n)$ en $\mathbb{F}_2^{\mathbb{F}_2^n}$, llamado de las *funciones lineales*, isomorfo a $\mathbb{F}_2^n$. De hecho, para cada $f \in \mathrm{Lin}(\mathbb{F}_2^n)$ existe un único $\boldsymbol{\delta} \in \mathbb{F}_2^n$ tal que $f(\boldsymbol{\varepsilon}) = \langle \boldsymbol{\delta} | \boldsymbol{\varepsilon} \rangle = \sum_{j=0}^{n-1} \delta_j \varepsilon_j$, para toda $\boldsymbol{\varepsilon} \in \mathbb{F}_2^n$. Escribiremos $f = \lambda_{\boldsymbol{\delta}}$.

El *complemento* de una función booleana $f \in \mathbb{F}_2^{\mathbb{F}_2^n}$ es $\overline{f} = 1 + f : \boldsymbol{\varepsilon} \mapsto 1 + f(\boldsymbol{\varepsilon})$.
Los complementos de las funciones lineales son las funciones *afines*: $\mathrm{Afín}(\mathbb{F}_2^n) = \{\overline{f} | \, f \in \mathrm{Lin}(\mathbb{F}_2^n)\}$.
Recordamos que el producto en $\mathbb{F}_2$ es idempotente, $x^2 = x$, y la suma es de orden 2: $x + x = 0$.

## 7.2 Formas algebraicas

Sean $X_0, \ldots, X_{n-1}$ $n$ símbolos de variables. Estas son pues entes sintácticos. A cada variable $X_j$ se le asocia con la proyección $\pi_j$. Así la connotación de la variable $X_j$ es el valor de la $j$-ésima entrada en cada punto de $\mathbb{F}_2^n$. Los *monomios* son productos de variables (distintas a pares) y los *polinomios* son combinaciones lineales de monomios. El *grado* de un monomio es el número de variables, distintas a pares, que aparecen en él como factores, y el *grado* de un polinomio es el mayor de los grados de los monomios que aparecen en él como sumandos. Un polinomio $P(X_0, \ldots, X_{n-1}) \in \mathbb{F}_2[X_0, \ldots, X_{n-1}]$ define la función $\mathbb{F}_2^n \to \mathbb{F}_2$, $\boldsymbol{\varepsilon} \mapsto P(\varepsilon_0, \ldots, \varepsilon_{n-1})$.

**Definición 7.2** *Para cada función booleana $f \in \mathbb{F}_2^{\mathbb{F}_2^n}$, su* soporte *es* $\mathrm{Spt}(f) = f^{-1}(1) = \{\boldsymbol{\varepsilon} \in \mathbb{F}_2^n | \, f(\boldsymbol{\varepsilon}) = 1\}$ *y su* parte nula *es* $\mathrm{Nul}(f) = f^{-1}(0) = \{\boldsymbol{\varepsilon} \in \mathbb{F}_2^n | \, f(\boldsymbol{\varepsilon}) = 0\}$. *Así,* $\{\mathrm{Nul}(f), \mathrm{Spt}(f)\}$ *es una partición de* $\mathbb{F}_2^n$.

**Definición 7.3** *Las siguientes nociones son convencionales:*

- *Para cada variable $X_j$, se hace $X_j^1 = X_j$ y $X_j^0 = 1$.*
- *Para cada $\boldsymbol{\varepsilon} = (\varepsilon_0, \ldots, \varepsilon_{n-1}) \in \mathbb{F}_2^n$ su* monomio característico *es $\mu(\boldsymbol{\varepsilon}) = \prod_{j=0}^{n-1} X_j^{\varepsilon_j}$.*

**Observación 7.2** *Las siguientes aseveraciones se cumplen para cada $n \in \mathbb{N}$:*

- *Toda función booleana es equivalente a un polinomio. O si se quiere, toda palabra-$2^n$ es equivalente a un polinomio.*
- *Toda función es lineal o afín cuando y sólo cuando sea de grado a lo sumo 1.*
- *El máximo grado posible es $n$.*

En efecto, puede verse que si $f \in \mathbb{F}_2^{\mathbb{F}_2^n}$ entonces ella coincide con la función

$$A_{fn} : (\mathbf{x}, x_{n-1}) \mapsto A_{fn}(\mathbf{x}, x_{n-1}) = (1 + x_{n-1}) f(\mathbf{x}, 0) + x_{n-1} f(\mathbf{x}, 1). \tag{14}$$

Al representar a las funciones $\mathbb{F}_2^{n-1} \to \mathbb{F}_2$, $\mathbf{x} \mapsto f(\mathbf{x}, 0)$, $\mathbf{x} \mapsto f(\mathbf{x}, 1)$ también como polinomios, de (14) se encuentra el polinomio equivalente a $f$. □

**Definición 7.4** *El polinomio equivalente a $f \in \mathbb{F}_2^{\mathbb{F}_2^n}$ se dice ser la* forma normal algebraica $\mathrm{FNA}(f)$ *de $f$.*

Una manera alternativa de calcular $\mathrm{FNA}(f)$ es la siguiente: Inicialmente hágase $g(X_0, \ldots, X_{n-1}) = 0$, y luego, para cada $\boldsymbol{\varepsilon} \in \mathbb{F}_2^n$, si acaso $g(\boldsymbol{\varepsilon}) \neq f(\boldsymbol{\varepsilon})$ actualícese $g(X_0, \ldots, X_{n-1}) := g(X_0, \ldots, X_{n-1}) + \mu(\boldsymbol{\varepsilon})$ donde $\mu(\boldsymbol{\varepsilon})$ es el monomio característico de $\boldsymbol{\varepsilon}$ (véase la definición 7.3). □

Ahora bien, al ser $\mathbb{F}_2$ un campo, se tiene, por el Teorema Fundamental del Algebra, que cualquier polinomio en $\mathbb{F}_2[X]$ que sea no-nulo a lo más posee un número de raíces igual a su grado. De aquí se sigue:

**Observación 7.3** *La colección $\{\mu(\boldsymbol{\varepsilon})\}_{\boldsymbol{\varepsilon} \in \mathbb{F}_2^n}$ es linealmente independiente y por tanto es una base de $\mathbb{F}_2^{\mathbb{F}_2^n}$ sobre $\mathbb{F}_2$.*

Observamos también que si $\boldsymbol{\varepsilon}_0 \in \mathbb{F}_2^n$ tiene un peso de Hamming $k = w(\boldsymbol{\varepsilon}_0)$ y es precisamente en los índices $j_1, \ldots, j_k$ que $\varepsilon_{0j_\kappa} = 1$ entonces para cualquier $\boldsymbol{\varepsilon}_1 \in \mathbb{F}_2^n$ se tiene

$$\mu(\boldsymbol{\varepsilon}_0)(\boldsymbol{\varepsilon}_1) = 1 \iff \forall \kappa \in [\![1, k]\!] : \varepsilon_{0j_\kappa} = 1.$$

Es decir $\mathrm{Spt}\,(\mu(\boldsymbol{\varepsilon}_0)) = \prod_{j=0}^{n-1} F_j$ donde $F_j = \{1\}$ si $\exists \kappa \in [\![1, k]\!]$: $j = j_\kappa$, o $F_j = \{0, 1\}$ en otro caso.

**Observación 7.4** *El soporte de cada monomio $\mu(\boldsymbol{\varepsilon})$ es una variedad lineal de dimensión $n - w(\boldsymbol{\varepsilon})$ en $\mathbb{F}_2^n$. Posee, por tanto, $2^{n - w(\boldsymbol{\varepsilon})}$ elementos.*

## 7.3 Códigos de Reed-Muller

Sea $\mathbb{F}_2[X_0,\ldots,X_{n-1}]^{\leq r}$ la colección de polinomios con grado a lo sumo $r$, el cual es un subespacio lineal de $\mathbb{F}_2[X_0,\ldots,X_{n-1}]$. Por la observación 7.3, una base de ella es $\{\mu(\boldsymbol{\varepsilon})\}_{w(\boldsymbol{\varepsilon})\leq r}$, y es, por tanto, de dimensión

$$k_{nr} = \sum_{j=0}^{r} \binom{n}{j}. \tag{15}$$

**Definición 7.5 (Reed-Muller)** *Para $m, r \in \mathbb{N}$, sea*

$$\mathcal{R}(m,r) = \left\{ [f(\boldsymbol{\varepsilon})]_{\boldsymbol{\varepsilon}\in\mathbb{F}_2^m} \in \mathbb{F}_2^{2^m} \mid f \in \mathbb{F}_2[X_0,\ldots,X_{m-1}]^{\leq r} \right\}$$

*la colección de palabras de código resultantes como evaluaciones de los polinomios de grado a lo sumo $r$ en el espacio $\mathbb{F}_2^m$ de dimensión $m$.*

Se tiene, naturalmente, que $\mathcal{R}(m,r)$ es un código-$[2^m, k_{mr}]$, donde $k_{mr}$ está dado por la ec. (15). Así pues sus palabras de código son de longitud $2^m$ y poseen $k_{mr}$ bits de información. Una generatriz de él es la matriz de orden $2^m \times k_{mr}$:

$$R_{mr} = \left[\mu(\boldsymbol{\delta}_j)(\boldsymbol{\varepsilon}_i)\right]_{(i,j)\in[\![0,2^m-1]\!]\times[\![0,k_{mr}-1]\!]},$$

donde $\{\boldsymbol{\delta}_j\}_{j\in[\![0,k_{mr}-1]\!]}$ es una enumeración de los puntos en $\mathbb{F}_2^m$ con peso de Hamming a lo sumo $r$.

Por ejemplo, para $m = 4$, el código $\mathcal{R}(m, m-1)$ posee como generatriz a

$$R_{43} = \left(\begin{array}{c|cccc|cccccc|cccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right)$$

La primera columna contiene los valores, en $\mathbb{F}_2^4$, del único polinomio de grado 0 a saber la constante 1, las siguientes 4 los de las variables $X_i$, $i \in [\![0,3]\!]$, las siguientes 6 los de las cuadráticas $X_iX_j$, $\{i,j\} \in [\![0,3]\!]^{(2)}$, y las últimas 4 los de las cúbicas $X_iX_jX_k$, $\{i,j,k\} \in [\![0,3]\!]^{(3)}$. La primera columna contiene $16 = 2^{4-0}$ 1's, cada una de las 4 siguientes $8 = 2^{4-1}$, cada una de las 6 siguientes $4 = 2^{4-2}$, y cada una de las 4 últimas $2 = 2^{4-3}$.

De la observación 7.4 se sigue:

**Observación 7.5** *El código $\mathcal{R}(m,r)$ tiene peso mínimo $2^{m-r}$.*

**Proposición 7.1** *El dual del código $\mathcal{R}(m,r)$ es el código de Reed-Muller $\mathcal{R}(m, m-r-1)$, es decir*

$$\mathcal{R}(m,r)^\perp = \mathcal{R}(m, m-r-1).$$

En efecto, primero se tiene que la dimensión del dual $\mathcal{R}(m,r)^\perp$ es

$$2^m - k_{mr} = \sum_{j=0}^{m}\binom{m}{j} - \sum_{j=0}^{r}\binom{m}{j} = \sum_{j=r+1}^{m}\binom{m}{j} = \sum_{j=0}^{m-r-1}\binom{m}{j} = k_{m,m-r-1}.$$

Así, al ver que $\mathcal{R}(m, m-r-1) \subset \mathcal{R}(m,r)^\perp$ se tendrá que estos espacios coinciden. Para ello si se toma a dos palabras en los códigos $\boldsymbol{\varepsilon} \in \mathcal{R}(m,r)$ y $\boldsymbol{\delta} \in \mathcal{R}(m, m-r-1)$, entonces han de existir dos polinomios $f, g \in \mathbb{F}_2[X_0, \ldots, X_{m-1}]$ de grados respectivos $e \in [\![0, r]\!]$, $d \in [\![0, m-r-1]\!]$, que producen $\boldsymbol{\varepsilon}$ y $\boldsymbol{\delta}$ al evaluarlos en $\mathbb{F}_2^m$. El producto $h = fg$ es un polinomio de grado a lo sumo $e + d \leq r + (m-r-1) = m-1$. Por tanto $[h(\boldsymbol{\eta})]_{\boldsymbol{\eta} \in \mathbb{F}_2^m}$ es una palabra en el código $\mathcal{R}(m, m-1)$ el cual consiste sólo de palabras de peso de Hamming par. Con esto resulta que necesariamente $\langle \boldsymbol{\varepsilon} | \boldsymbol{\delta} \rangle = 0$, es decir, esas dos palabras de código son ortogonales. □

**Definición 7.6** *Para cualquier conjunto $A \subset \mathbb{F}_2^m$ su* función característica *es $\xi_A : \boldsymbol{\varepsilon} \mapsto \xi_A(\boldsymbol{\varepsilon})$ donde*

$$\xi_A(\boldsymbol{\varepsilon}) = \begin{cases} 1 & si \;\; \boldsymbol{\varepsilon} \in A \\ 0 & si \;\; \boldsymbol{\varepsilon} \notin A \end{cases}$$

*Así, $\xi_A \in \mathbb{F}_2^{\mathbb{F}_2^m}$.*

**Observación 7.6** *Sea $\boldsymbol{\delta} \in \mathbb{F}_2^n$ un punto en el hipercubo y sea $J_0 = \boldsymbol{\delta}^{-1} \subset [\![0, n-1]\!]$ el conjunto de índices con entradas 1 en $\boldsymbol{\delta}$. Entonces*

$$\forall \boldsymbol{\varepsilon} \in \mathbb{F}_2^n : \quad \xi_{\{\boldsymbol{\delta}\}}(\boldsymbol{\varepsilon}) = \prod_{j=0}^{n-1} (1 + \delta_j + \varepsilon_j) = \sum_{J_0 \subset J \subset [\![0,n-1]\!]} \prod_{j \in J} \varepsilon_j. \tag{16}$$

*Así pues, la función característica de la mónada $\{\boldsymbol{\delta}\}$ está dada por el polinomio $\sum \{\mu(\boldsymbol{\eta}) |\, \boldsymbol{\delta} \preceq \boldsymbol{\eta}\}$, donde $\preceq$ es el orden del hipercubo.*

Se tiene entonces que para cualquier función booleana $f \in \mathbb{F}_2^{\mathbb{F}_2^n}$:

$$\begin{aligned} f &= \sum \left\{ f(\boldsymbol{\delta}) \xi_{\{\boldsymbol{\delta}\}} |\, \boldsymbol{\delta} \in \mathbb{F}_2^n \right\} \\ &= \sum \left\{ f(\boldsymbol{\delta}) \sum \{\mu(\boldsymbol{\eta}) |\, \boldsymbol{\delta} \preceq \boldsymbol{\eta}\} \,|\, \boldsymbol{\delta} \in \mathbb{F}_2^n \right\} \\ &= \sum \left\{ \left( \sum \{f(\boldsymbol{\delta}) |\, \boldsymbol{\delta} \preceq \boldsymbol{\eta}\} \right) \mu(\boldsymbol{\eta}) |\, \boldsymbol{\eta} \in \mathbb{F}_2^n \right\} \\ &= \sum \{ g_f(\boldsymbol{\eta}) \mu(\boldsymbol{\eta}) |\, \boldsymbol{\eta} \in \mathbb{F}_2^n \} \end{aligned}$$

donde $g_f(\boldsymbol{\eta}) = \sum \{f(\boldsymbol{\delta}) |\, \boldsymbol{\delta} \preceq \boldsymbol{\eta}\}$.

De manera más precisa:

**Observación 7.7** *Si $A = \boldsymbol{\varepsilon} + V$ es un conjunto* afín*, es decir $V$ es un espacio lineal en $\mathbb{F}_2^m$, de dimensión $r$, entonces existe un polinomio $f_A \in \mathbb{F}_2[X_0, \ldots, X_{m-1}]$ de grado $m-r$ tal que $f_A = \xi_A$.*

En efecto, $A$ puede ser visto como el conjunto de soluciones de un sistema lineal $A\mathbf{x} = \mathbf{b}$ de $m$ incógnitas y $m-r$ ecuaciones, que puede reescribirse como uno de la forma $A\mathbf{x} + \mathbf{b} + \mathbf{1} = \mathbf{1}$. Entonces

$$\xi_A(\mathbf{x}) = \prod_{i=0}^{m-r-1} \sum_{j=0}^{m-1} (a_{ij} x_j + b_i + 1) \,.$$

Esta última expresión determina al polinomio $f_A$. □

De aquí se siguen sin más:

**Observación 7.8** *El código de Reed-Muller $\mathcal{R}(m,r)$ es el espacio generado por las funciones características de las variedades afines de dimensión al menos $m-r$:*

$$\mathcal{R}(m,r) = \mathcal{L}\left\{ \xi_A |\; A = \boldsymbol{\varepsilon} + V \,,\, \dim V \geq m-r \,,\, \boldsymbol{\varepsilon} \in \mathbb{F}_2^m \right\}.$$

**Observación 7.9** *La función característica de cualquier variedad afín de dimensión $r+1$ está en el código dual de $\mathcal{R}(m,r)$. Por tanto,*

$$A = \boldsymbol{\varepsilon} + V \,,\, \dim V = r+1 \implies \xi_A^T R_{mr} = \mathbf{0}. \tag{17}$$

## 7.4 Decodificación de códigos de Reed-Muller

Veremos algunas ideas expuestas en [2] y en su actualización [3].

Se tiene que el código de Reed-Muller $\mathcal{R}(m,r)$ es un código-$[2^m, k_{mr}, 2^{m-r}]$, por tanto, de acuerdo con la observación 4.2, se tiene que puede corregirse con él menos de $\frac{2^{m-r}}{2} = 2^{m-r-1}$ errores.

Veamos cómo, habiéndose recibido una palabra $\boldsymbol{\delta} = (\delta_i)_{i=0}^{2^m-1} \in \mathbb{F}_2^{2^m}$, cuando se quería transmitir $\boldsymbol{\varepsilon} = (\varepsilon_i)_{i=0}^{2^m-1} \in \mathcal{R}(m,r)$, se corrige hasta $2^{m-r-1}-1$ bits en ella para recuperar la palabra $\boldsymbol{\varepsilon} \in \mathcal{R}(m,r)$ en el código más cercana a $\boldsymbol{\delta}$.

Para $i \in [\![0, 2^m-1]\!]$ sea $(i)_2 \in \mathbb{F}_2^m$ el vector que coincide con la representación en base-2 de $i$, con $n$ bits, el menos significativo hacia la derecha, y sea $\{(i)_2\} = (i)_2 + \{\mathbf{0}\}$ la mónada que consiste del vector $(i)_2$. Claramente, $\{(i)_2\}$ es una variedad afín de dimensión cero en el hipercubo $\mathbb{F}_2^m$.

Para cada variedad afín $A$ de dimensión $n \in [\![0, r+1]\!]$ en $\mathbb{F}_2^m$, digamos que ésta es *par* o *impar* según lo sea $\mathrm{card}\{i \in [\![0, 2^m-1]\!] \mid (i)_2 \in A \,\&\, \delta_i \neq \varepsilon_i\}$, es decir, según sea la paridad del número de "errores" en $A$.

Pues bien, si $A = \boldsymbol{\varepsilon} + V$, con $\dim V = r+1$, es una variedad afín de dimensión $r+1$, entonces de acuerdo con la relación (17) si $\boldsymbol{\delta}$ estuviese en el código, entonces $\langle \xi_A | \boldsymbol{\delta} \rangle = 0$. Así pues se tiene que la paridad de $A$ es necesariamente $\langle \xi_A | \boldsymbol{\delta} \rangle$. Consecutivamente, si $A = \boldsymbol{\varepsilon} + V$, con $\dim V = n \leq r$, es una variedad afín de dimensión $n$, entonces se decidirá si es par o impar por *mayoría de votos*: Se ha de tener que $A$ está incluída en un número impar de variedades afines de dimensión $n+1$, algunas pares y otras impares. La paridad de $A$ será aquella que resulte mayoritaria entre estas últimas.

**Proposición 7.2** *Si $A$ es una variedad afín de dimensión $n$ en $\mathbb{F}_2^m$, entonces está incluída en exactamente $2^{m-n}-1$ variedades afines de dimensión $n+1$. De hecho cada uno de los puntos en el complemento de $A$ pertenece a una única de tales variedades afines.*

En efecto, supongamos $A = \boldsymbol{\varepsilon} + V$, con $\dim V = n$. Veremos primero que el subsepacio $V$ está en $2^{m-n}-1$ subespacios de dimensión $n+1$. Si $\boldsymbol{\varepsilon}_1 \notin V$ entonces $V \oplus \boldsymbol{\varepsilon}_1 = \{\mathbf{v} + t\boldsymbol{\varepsilon}_1 \mid \mathbf{v} \in V\,,\, t \in \mathbb{F}_2\}$ es un subespacio de dimensión $n+1$ que contiene a $V$. Ahora bien, se tiene $V \oplus \boldsymbol{\varepsilon}_1 = V \oplus \boldsymbol{\varepsilon}_2$ si y sólo si $\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2 \in V$. Así pues, el número de espacios distintos de la forma $V \oplus \boldsymbol{\varepsilon}$ coincide con el de clases laterales que define $V$.

Ya que la cardinalidad del espacio cociente $\mathbb{F}_2^m/V$ es $\frac{2^m}{2^n} = 2^{m-n}$, se tiene que el número de subespacios de dimensión $n+1$ que contienen a $V$ es $2^{m-n}-1$ (ésos son de la forma $V \oplus \boldsymbol{\varepsilon}$ con $\boldsymbol{\varepsilon} \neq \mathbf{0}$).

Ahora bien, $A' = \boldsymbol{\varepsilon} + V'$ es una extensión de dimensión $n+1$ de $A = \boldsymbol{\varepsilon} + V$ si y sólo si $V'$ es una extensión de dimensión $n+1$ de $V$. Por tanto hay $2^{m-n}-1$ de tales extensiones. □

**Proposición 7.3** *Si $\boldsymbol{\delta} = (\delta_i)_{i=0}^{2^m-1} \in \mathbb{F}_2^{2^m}$ involucra menos de $2^{m-r-1}$ errores, entonces siempre que $A$ sea una variedad afín de dimensión $n \leq r$ en $\mathbb{F}_2^m$, su paridad coincidirá con la paridad mayoritaria entre las variedades de dimensión $n+1$ que contengan a $A$.*

En efecto, supongamos que se haya cometido $t < 2^{m-r-1}$ errores. Sea $A$ una variedad afín de dimensión $r$ en $\mathbb{F}_2^m$ y sean $(i_1)_2, \ldots (i_s)_2$ los errores cometidos fuera de $A$. Por un lado $2t < 2^{m-r}-1$ y en consecuencia $t < (2^{m-r}-1) - t$. El lado izquierdo es una cota superior de $s$, que cuenta el número de errores fuera de $A$. El lado derecho, en cambio, cuenta el número de espacios de dimensión $r$ en donde no hay error. Por tanto todos estos espacios han de tener la misma paridad que $A$, y ellos son más.

Este argumento se lleva a dimensiones menores. □

**Procedimiento de decodificación** Habiendo recibido $\boldsymbol{\delta} = (\delta_i)_{i=0}^{2^m-1} \in \mathbb{F}_2^{2^m}$:

- ***Paso inicial.*** Para cada variedad afín $A$ de dimensión $r+1$ declárese par o impar según sea el valor $\langle \xi_A | \boldsymbol{\delta} \rangle$.
- ***Paso recursivo.*** Para cada $n = r, \ldots, 1, 0$ declárese a cada variedad afín $A$ de dimensión $n$ según haya sido declarada la mayoría de las variedades de dimensión $n+1$ que contengan a $A$.
- ***Paso final.*** En las variedades de dimensión 0, aquellas que sean impares dan los índices en donde hay que cambiar el valor de $\boldsymbol{\delta}$.

El número de conjuntos linealmente independientes de cardinalidad $k$ en $\mathbb{F}_2^n$ es $\alpha_{nk} = \prod_{\kappa=0}^{k-1}(2^n - 2^\kappa)$. Evidentemente, el número de bases de un espacio de dimensión $k$ es $\alpha_{kk}$. Así, el número de espacios de dimensión $k$ en $\mathbb{F}_2^n$ es

$$\beta_{nk} = \frac{\alpha_{nk}}{\alpha_{kk}} = \prod_{\kappa=0}^{k-1} \frac{2^n - 2^\kappa}{2^k - 2^\kappa} = \prod_{\kappa=0}^{k-1} \frac{2^{n-\kappa} - 1}{2^{k-\kappa} - 1} = \prod_{\kappa=1}^{k} \frac{2^{n-k+\kappa} - 1}{2^\kappa - 1}.$$

Ya que cada espacio de dimensión $k$ determina $2^{n-k}$ variedades afines paralelas a él, se tiene finalmente que el número de variedades afines de dimensión $k$ en $\mathbb{F}_2^n$ es

$$2^{n-k}\beta_{nk} = 2^n \prod_{\kappa=1}^{k} \frac{2^{n-k+\kappa} - 1}{2(2^\kappa - 1)}. \tag{18}$$

De hecho este método de recuento de variedades afines sirve de base para diseñar uno que permita enumerar, una a una, las variedades afines de dimensión $r+1$ requeridas en el "Paso inicial" del algoritmo de decodificación de Reed-Muller.

El "Paso recursivo" se puede realizar enumerando a las variedades de dimensión mayor en uno que la actual y que la contienen utilizando la proposición 7.2.

# 8 Códigos cíclicos

## 8.1 Polinomios generadores y revisores de paridad

**Definición 8.1** *Sea $\mathbb{K}$ un campo finito. En el espacio $\mathbb{K}^n$, la transformación lineal que sobre la base canónica actúa como $\rho_n : \mathbf{e}_j \mapsto \mathbf{e}_{(j+1) \bmod n}$ se dice ser la* rotación de componentes. *Un código lineal-$(n,k)$ $C < \mathbb{K}^n$ es* cíclico *si $\rho_n(C) \subset C$.*

Sea $\mathbb{K}[X]$ el anillo de polinomios sobre $\mathbb{K}$ y sea $\iota_n : \mathbb{K}^n \to \mathbb{K}[X]$, $(a_0, \ldots, a_{n-1}) \mapsto \sum_{i=0}^{n-1} a_i X^i$ la transformación que a cada vector de $n$ coeficientes lo convierte en el polinomio correspondiente. $\iota_n$ es, naturalmente, un monomorfismo lineal de espacios vectoriales sobre $\mathbb{K}$. Consideremos el polinomio $c_n(X) = X^n - 1$ y al cociente $\mathbb{K}[X]/(c_n(X)) = \mathbb{K}[X]/(X^n - 1)$ el cual, visto como un espacio vectorial de dimensión $n$, hace que $\pi \circ \iota_n : \mathbf{a} \mapsto \iota_n(\mathbf{a}) + (X^n - 1)$ sea un isomorfismo lineal que aplica la base canónica de $\mathbb{K}^n$ sobre la *base polinomial* $\big(X^j + (X^n - 1)\big)_{j=0}^{n-1}$, donde $\pi : \mathbb{K}[X] \to \mathbb{K}[X]/(X^n - 1)$ es la *proyección canónica* $\pi : f(X) \mapsto f(X) + (X^n - 1)$. Se tiene el diagrama:

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{\iota_n} & \mathbb{K}[X] \\ & \searrow^{\pi_n \circ \iota_n} & \downarrow{\scriptstyle \pi_n} \\ & & \mathbb{K}[X]/(X^n - 1) \end{array} \tag{19}$$

Si $C < \mathbb{K}^n$ es un código cíclico entonces vale la implicación:

$$\mathbf{a} \in C \implies X \cdot (\pi \circ \iota_n(\mathbf{a})) \in \pi \circ \iota_n(C),$$

por tanto, la imagen $\pi \circ \iota_n(C)$ es un ideal en el anillo $\mathbb{K}[X]/(X^n - 1)$.

Como $\mathbb{K}[X]$ es un anillo de ideales principales, necesariamente existe un polinomio $g(X) \in \mathbb{K}[X]$ tal que $(g(X) + (X^n - 1)) = \pi \circ \iota_n(C)$. Un tal polinomio con grado mínimo se dice ser *generador* del código $C$. De hecho el grado mínimo ha de ser $n - k$ y en $\mathbb{K}[X]$ se ha de tener $g(X) | (X^n - 1)$.

Recíprocamente, si $g(X) \in \mathbb{K}[X]$ es tal que $g(X) | (X^n - 1)$, el ideal generado por él, reducido módulo $(X^n - 1)$, consta de polinomios cuyos vectores de coeficientes forman un código cíclico $C$.

**Observación 8.1** *Así para cada $n$ por cada divisor del polinomio $X^n - 1$ en $\mathbb{K}[X]$ se tendrá un código cíclico.*

Hagamos aquí una breve disgresión sobre divisores del polinomio $X^n - 1$ en campos finitos $\mathbb{K} = \mathbb{F}_q$, donde $q$ es la potencia de un primo. El *orden* de un polinomio no-nulo $p(X) \in \mathbb{F}_q[X]$ es el mínimo $n$ tal que $p(X)|(X^n - 1)$ en $\mathbb{F}_q[X]$. Un polinomio $p(X) \in \mathbb{F}_q[X]$ de grado $m \in \mathbb{N}$ es *primitivo* si es el polinomio mínimo de un elemento primitivo, es decir, de un generador del grupo multiplicativo $\mathbb{F}_{q^m}^*$. Pues bien, se tiene que un polinomio de grado $m$ es primitivo cuando y sólo cuando su orden es $q^m - 1$. Además, para cada $m$ el número de polinomios primitivos de grado $m$ en $\mathbb{F}_q[X]$ es $\phi(q^m - 1)/m$ donde $\phi$ es la función tociente de Euler.

Así para un polinomio primitivo $g(X) \in \mathbb{F}_2[X]$ de grado $m$, de acuerdo con la observación 8.1, el tamaño de bloque para que $g(X)$ sea el generador de un código cíclico debe ser $n = 2^m - 1$.

Sigamos con nuestra exposición. Si el polinomio generador de un código cíclico es $g(X) = \sum_{i=0}^{n-k} g_i X^i$ entonces

$$G_{nk} = \begin{pmatrix} g_0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ g_1 & g_0 & 0 & \cdots & 0 & 0 & 0 \\ g_2 & g_1 & g_0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ g_{n-k} & g_{n-k-1} & g_{n-k-2} & \cdots & 0 & 0 & 0 \\ 0 & g_{n-k} & g_{n-k-1} & \cdots & 0 & 0 & 0 \\ 0 & 0 & g_{n-k} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & g_{n-k} & g_{n-k-1} & g_{n-k-2} \\ 0 & 0 & 0 & \cdots & 0 & g_{n-k} & g_{n-k-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & g_{n-k} \end{pmatrix} \in \mathbb{K}^{n \times k} \tag{20}$$

es una generatriz del código $C$. El polinomio generador $g(X)$, que es único si se le supone *mónico*, es decir con coeficiente principal 1, determina pues por completo al código cíclico $C$. Es convencional identificar a cada palabra de código $\mathbf{a} \in C$ tanto con el polinomio $\iota_n(\mathbf{a})$ como con la clase $\iota_n(\mathbf{a}) + (X^n - 1)$ de ese polinomio y por tanto en el contexto de códigos cíclicos se dice que *las palabras de código son polinomios*.

**Ejemplo 8.1 (Palabras de peso par)** *Sea $\mathbb{K} = \mathbb{F}_2$ el campo primo de característica 2 y sea $C \subset \mathbb{F}_2^n$ el espacio de palabras en $\mathbb{F}_2^n$ con peso de Hamming par.*

Este es un subespacio lineal y una matriz revisora de paridad es $\mathbf{1}^T$:

$$\forall \boldsymbol{\varepsilon} \in \mathbb{F}_2^n : \ \boldsymbol{\varepsilon} \in C \iff 0 = \mathbf{1}^T \boldsymbol{\varepsilon} = \langle \mathbf{1} | \boldsymbol{\varepsilon} \rangle.$$

Por tanto $n - k = 1$ y consecuentemente $C$ posee $k = n - 1$ bits de información.

Claramente $C$ es cíclico. Como la palabra $110^{n-2}$ está en $C$, el polinomio generador ha de ser $g(X) = X+1$ el cual evidentemente divide a $c_n(X) = X^n - 1 = X^n + 1$ en $\mathbb{F}_2[X]$. Algunos ejemplos de generatrices son

$$G_3 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \ , \ G_4 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \ , \ G_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ , \ G_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Así el código de palabras con peso par en $\mathbb{F}_2^n$ puede representarse de manera precisa con el solo polinomio $g(X) = X + 1$. □

Sea $C \subset \mathbb{K}^n$ un código cíclico y sea $g(X) = \sum_{i=0}^{n-k-1} g_i X^i + X^{n-k}$ su polinomio generador. Como $g(X)|(X^n - 1)$, el cociente $h(X) = \frac{X^n - 1}{g(X)}$ es un polinomio, de grado $k$. Escribamos $h(X) = \sum_{i=0}^{k} h_i X^i$. Entonces,

$$X^n - 1 = g(X)h(X) = \left( \sum_{i=0}^{n-k} g_i X^i \right) \left( \sum_{j=0}^{k} h_j X^j \right) = \sum_{\ell=0}^{n} \left( \sum_{i+j=\ell} g_i h_j \right) X^\ell$$

por lo cual en el campo $\mathbb{K}$ se ha de tener

$$g_0 h_0 = 1 \;\;\&\;\; \left[ \ell \in [\![1, n-1]\!] \;\Rightarrow \sum_{i=\max\{\ell-(n-k),0\}}^{\min\{\ell,k\}} g_i h_{\ell-i} = 0 \right] \;\;\&\;\; g_{n-k} h_k = 1. \tag{21}$$

Al considerar un polinomio $f(X) \in (g(X))$, necesariamente

$$h(X)\, f(X) = \frac{X^n - 1}{g(X)}\, e(X)\, g(X) = e(X)\,(X^n - 1)$$

para algún $e(X) \in \mathbb{K}[X]$, o sea $h(X)\, f(X) \equiv 0 \bmod (X^n - 1)$. Es debido a esto último que el polinomio $h(X)$ se dice ser el *polinomio revisor de paridad* del código cíclico $C$. De hecho la matriz

$$H_{nk} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & h_k & h_{k-1} & \cdots & h_2 & h_1 & h_0 \\ 0 & 0 & 0 & \cdots & h_k & h_{k-1} & h_{k-2} & \cdots & h_1 & h_0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & h_k & \cdots & h_3 & h_2 & h_1 & \cdots & 0 & 0 & 0 \\ 0 & h_k & h_{k-1} & \cdots & h_2 & h_1 & h_0 & \cdots & 0 & 0 & 0 \\ h_k & h_{k-1} & h_{k-2} & \cdots & h_1 & h_0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \in \mathbb{K}^{(n-k)\times n} \tag{22}$$

es revisora de paridad del código $C$, pues por las relaciones (21) se ve que $H_{nk} G_{nk} = 0 \in \mathbb{K}^{(n-k)\times k}$, es decir cada rengón de $H_{nk}$ es ortogonal a todas las columnas de $G_{nk}$. Es importante remarcar que en la ec. (20) los índices son crecientes en cada columna, en tanto que en la ec. (22) son decrecientes en cada renglón. Se tiene pues:

$$\forall \mathbf{u} \in \mathbb{K}^n :\; [\mathbf{u} \in C \implies H_{nk}\mathbf{u} = \mathbf{0}]\,,$$

y en consecuencia

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{K}^n :\; [\mathbf{u} - \mathbf{v} \in C \implies H_{nk}\mathbf{u} = H_{nk}\mathbf{v}]\,. \tag{23}$$

**Ejemplo 8.2 (Códigos cíclicos binarios de longitud $n = 7$)** *En $\mathbb{K} = \mathbb{F}_2$, el polinomio $X^7 - 1 = X^7 + 1$ se factoriza como $X^7 + 1 = (X+1)(X^3 + X + 1)(X^3 + X^2 + 1)$, así posee tres divisores irreducibles*

$$\gamma_{70}(X) = X + 1 \qquad , \qquad \gamma_{71}(X) = X^3 + X + 1 \qquad , \qquad \gamma_{72}(X) = X^3 + X^2 + 1.$$

*Cualquier producto de éstos dividirá a $X^7 + 1$, por lo que existen $2^3 = 8$ divisores de él, de los cuales dos son triviales: $g_{70}(X) = \gamma_{70}(X)\gamma_{71}(X)\gamma_{72}(X) = X^7 + 1$ y $g_{77}(X) = 1$.*

Hay pues $6 = 2^3 - 2$ códigos cíclicos binarios de longitud $n = 7$. En particular, para el polinomio $g_{73}(X) = \gamma_{70}(X)\gamma_{71}(X) = X^4 + X^3 + X^2 + 1$ se tiene $h_{73}(X) = \frac{X^7+1}{g_{73}(X)} = \gamma_{72}(X) = X^3 + X^2 + 1$ y las matrices generatriz y revisora de paridad siguientes:

$$G_{73} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad , \qquad H_{73} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

La función de codificación es pues

$$(u_0, u_1, u_2) \mapsto (u_0\,,\, u_1\,,\, u_0 + u_2\,,\, u_0 + u_1\,,\, u_0 + u_1 + u_2\,,\, u_1 + u_2\,,\, u_2).$$

Este se dice ser el *código "símplex"*. En los $7 = 2^3 - 1$ renglones de la matriz generatriz $G_{73}$ aparecen las representaciones en binario de los números en el intervalo $[\![1, 2^3 - 1]\!]$, por tanto su código dual es uno de Hamming. □

**Ejemplo 8.3 (Códigos cíclicos terciarios de longitud $n = 8$)** *En $\mathbb{K} = \mathbb{F}_3$, el polinomio $X^8-1 = X^8+2$ (pues $2 = -1$) se factoriza como $X^8 - 1 = (X + 1)(X + 2)(X^2 + 1)(X^2 + X + 2)(X^2 + 2X + 2)$, así posee cinco divisores irreducibles*

$$\gamma_{80}(X) = X + 1 \quad , \quad \gamma_{81}(X) = X + 2$$
$$\gamma_{82}(X) = X^2 + 1$$
$$\gamma_{83}(X) = X^2 + X + 2 \quad , \quad \gamma_{84}(X) = X^2 + 2X + 2.$$

*Cualquier producto de éstos dividirá a $X^8 - 1$, por lo que existen $2^5 = 32$ divisores de él, de los cuales dos son triviales: $g_{80}(X) = X^8 - 1$ y $g_{88}(X) = 1$.*

Hay pues $30 = 2^5 - 2$ códigos cíclicos terciarios de longitud $n = 8$. En particular, para el polinomio $g_{85}(X) = \gamma_{81}(X)\gamma_{82}(X) = X^3 + 2X^2 + X + 2$ se tiene $h_{85}(X) = \frac{X^8-1}{g_{85}(X)} = \gamma_{80}(X)\gamma_{83}(X)\gamma_{84}(X) = X^5 + X^4 + X + 1$ y las matrices generatriz y revisora de paridad siguientes:

$$G_{85} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad , \quad H_{85} = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

□

## 8.2 Codificación y decodificación

Sea $C \subset \mathbb{K}^n$ un código-$[n, k]$ cíclico, sea $g(X) \in \mathbb{K}[X]$ su polinomio generador y sea $G_{nk}$ su matriz generatriz. Naturalmente, dada una palabra $\mathbf{a} \in \mathbb{K}^k$, ella ha de quedar codificada por la palabra $\mathbf{c} = G_{nk}\mathbf{a} \in \mathbb{K}^n$. Ya que la $j$-ésima columna $g_j$ de la matriz generatriz $G_{nk}$ corresponde bajo la función $\iota_n$ al polinomio $X^j g(X)$, se tiene $\iota_n(\mathbf{c}) = \iota_n(G_{nk}\mathbf{a}) = g(X)\iota_k(\mathbf{a})$. En otras palabras, la manera de codificar a cada palabra es viéndola como un polinomio y multiplicando por éste al polinomio generador. Esta codificación mediante códigos cíclicos se dice ser *no-sistemática*.

**Definición 8.2** *Sea $C < \mathbb{K}^n$ un código cíclico con polinomio generador $g(X)$. Para cada $\mathbf{u} \in \mathbb{K}^n$, sea $\iota_n(\mathbf{u})$ el polinomio con vector de coeficientes $\mathbf{u}$. El* síndrome polinomial *de $\mathbf{u}$ es el polinomio $\iota_n(\mathbf{u}) \bmod g(X)$ (o si se quiere, su vector de coeficientes).*

Esta definición concuerda con la 5.5. En efecto, para cada $\mathbf{u} \in \mathbb{K}^n$, al escribir $r(X) = \iota_n(\mathbf{u}) \bmod g(X)$, entonces $\iota_n(\mathbf{u}) = q(X)\, g(X) + r(X)$, con grado$(r(X)) < n - k$, para algún polinomio $q(X)$. Así, sea $\mathbf{r} \in \mathbb{K}^n$ tal que $\iota_n(\mathbf{r}) = r(X)$ (como grado$(r(X)) < n - k$, las últimas $k$ entradas de $\mathbf{r}$ son 0). Por la relación (23),

$$H_{nk}\mathbf{u} = H_{nk}\mathbf{r} = \begin{bmatrix} H_{nk}^{(n-k)} & H_{nk}^{(k)} \end{bmatrix} \begin{bmatrix} \mathbf{r}^{(n-k)} \\ \mathbf{r}^{(k)} \end{bmatrix} = H_{nk}^{(n-k)}\mathbf{r}^{(n-k)}, \tag{24}$$

donde $H_{nk}^{(n-k)} \in \mathbb{K}^{(n-k)\times(n-k)}$ es la matriz formada por las primeras $(n - k)$ columnas de $H_{nk}$, $H_{nk}^{(k)} \in \mathbb{K}^{(n-k)\times k}$ es la matriz formada por las últimas $k$ columnas de $H_{nk}$, $\mathbf{r}^{(n-k)} \in \mathbb{K}^{n-k}$ consta de las primeras $(n - k)$ entradas de $\mathbf{r}$ y $\mathbf{r}^{(k)} = \mathbf{0} \in \mathbb{K}^k$ consta de las últimas $k$ entradas de $\mathbf{r}$. El síndrome de $\mathbf{u}$, en el sentido de la definición 5.5, es $H_{nk}\mathbf{u}$ y, según la relación (24), éste es $H_{nk}^{(n-k)}\mathbf{r}^{(n-k)}$. Ya que $H_{nk}^{(n-k)}$ posee una forma triangular su determinante es $(-1)^{\lfloor \frac{n-k}{2} \rfloor} h_k^{n-k}$ y es, por tanto, una matriz no singular. Esta determina un automorfismo lineal en $\mathbb{K}^{n-k}$ y la correspondencia entre los síndromes de la definición 5.5 y los de la 8.2.

**Procedimiento de decodificación.** Supóngase que al transmitir una palabra $\mathbf{a} \in C$ se recibiera la palabra $\mathbf{b} \in \mathbb{K}^n$. El error cometido sería pues $\mathbf{e} = \mathbf{a} - \mathbf{b}$. Vistas las palabras como polinomios, al calcular el síndrome polinomial $r(X) \in \mathbb{K}[X]$ de $\mathbf{b}$ se tendrá que existe un polinomio $q(X) \in \mathbb{K}[X]$ tal que $\iota_n(\mathbf{b}) = q(X)\, g(X) + r(X)$. Si la distancia mínima del código $C$ fuese $d$ entonces habría que localizar una palabra $\mathbf{e}_s$ de peso de Hamming a lo sumo $\left\lfloor \frac{d}{2} \right\rfloor$ tal que $\iota_n(\mathbf{e}_s) = r(X) \bmod g(X)$. Entonces necesariamente $\mathbf{e}_s = \mathbf{e}$ y en tal caso se corrige $\mathbf{b}$ cambiándolo por $\mathbf{a} = \mathbf{e}_s + \mathbf{b}$. □

El problema de cálculo de distancias mínimos de códigos cíclicos se ha tratado con diversos enfoques y el artículo de van Lint y Wilson [14] se ha convertido en una referencia clásica.

El proceso de decodificación descrito requiere pues calcular a los elementos con menor peso de Hamming en clases de congruencia módulo el polinomio generador.

**Observación 8.2** *Para un polinomio $f(X) \in \mathbb{K}[X]$ cualquiera, si $\overline{f}(X)$ es el polinomio que se obtiene al rotar los coeficientes de $f(X)$, entonces*

$$\overline{f}(X) \bmod g(X) = (X \cdot f(X)) \bmod g(X).$$

*Así pues, si $\iota_n(\mathbf{f}) = f(X)$, el síndrome polinomial del "polinomio rotado" $\iota_n(\rho_n(\mathbf{f}))$ coincide con el de $X\, r(X)$, donde $r(X)$ es el síndrome de $f(X)$.*

Observamos que si $f(X) = q(X)\, g(X) + r(X)$, con $f(X) = \iota_n(\mathbf{f})$, al rotar $\mathbf{f}$, valen:

$$\begin{aligned}
\iota_n(\rho_n(\mathbf{f})) &= X\, f(X) - f_{n-1}X^n + f_{n-1} \\
&= X\, f(X) - f_{n-1}(X^n - 1) \\
&= X\, f(X) - f_{n-1}\, g(X)h(X) \\
&= X(q(X)\, g(X) + r(X)) - f_{n-1}\, g(X)h(X) \\
&= X\, r(X) + g(X)(X\, q(X) - f_{n-1}h(X))
\end{aligned}$$

□

De esta manera, se puede calcular representantes principales para síndromes obtenidos de rotaciones de otros síndromes.

Por ejemplo, para el código-$[7, 3]$ símplex tratado en el ejemplo 8.2, el polinomio generador es $g_{73}(X) = X^4 + X^3 + X^2 + 1$. Por tanto, módulo $g_{73}(X)$, se tiene $X^4 = X^3 + X^2 + 1$, y en consecuencia

$$\begin{aligned}
X^5 &= X \cdot X^4 = X^4 + X^3 + X = \left(X^3 + X^2 + 1\right) + X^3 + X = X^2 + X + 1 \\
X^6 &= X \cdot X^5 = X^3 + X^2 + X \\
X^7 &= X \cdot X^6 = X^4 + X^3 + X^2 = \left(X^3 + X^2 + 1\right) + X^3 + X^2 = 1
\end{aligned}$$

Así pues, en la base polinomial se tiene la correspondencia siguiente:

| $\mathbf{e}$ | 1 | $X$ | $X^2$ | $X^3$ | $X^4$ | $X^5$ | $X^6$ |
|---|---|---|---|---|---|---|---|
| síndrome($\mathbf{e}$) | 1 | $X$ | $X^2$ | $X^3$ | $X^3 + X^2 + 1$ | $X^2 + X + 1$ | $X^3 + X^2 + X$ |

Dada una palabra $\mathbf{b}$, se calcula su síndrome, $r(X)$. Si éste aparece en el segundo renglón de la tabla anterior, entonces la correspondiente posición en el primero indicará cuál es el bit que hay que corregir. Si acaso el síndrome no apareciese, entonces, por tratarse de un código cíclico, la palabra puede rotarse y su síndrome multiplicarse por $X$ y reducirlo módulo $g(X)$ para volver a realizar la prueba. □

Otra manera de codificación, llamada ésta *sistemática*, utilizando códigos cíclicos es la siguiente. Dada una palabra $\mathbf{a} \in \mathbb{K}^k$ sea $\eta_{nk}(\mathbf{a}) = \sum_{\kappa=0}^{k-1} a_\kappa X^{n-k+\kappa}$ el polinomio de grado $n-1$ cuyos coeficientes "más altos" están dados por la palabra $\mathbf{a}$. Entonces ésta queda codificada por la palabra $\mathbf{c} \in \mathbb{K}^n$ tal que

$$\iota_n(\mathbf{c}) = \eta_{nk}(\mathbf{a}) - [\eta_{nk}(\mathbf{a}) \bmod g(X)]$$

o puesto equivalentemente,

$$c(X) = -\left[\left(X^{n-k}\, a(X)\right) \bmod g(X)\right] + X^{n-k}\, a(X)$$

el cual polinomio, en efecto, está en el código cíclico $C$ pues es un múltiplo del polinomio generador $g(X)$ de $C$. La palabra $\mathbf{c} \in \mathbb{K}^n$ de código, se descompone naturalmente en dos tramos: $\mathbf{c} = \mathbf{c}_r\mathbf{c}_i \in \mathbb{K}^{n-k} \times \mathbb{K}^k$, tales que la parte "alta" $\mathbf{c}_i \in \mathbb{K}^k$ coincide con la palabra *de información* $\mathbf{a}$ y la parte "baja" $\mathbf{c}_r \in \mathbb{K}^{n-k}$ tiene fines de *revisión de paridad*.

Ahora, si Alicia envíase una palabra de código $\mathbf{c} = \mathbf{c}_r\mathbf{c}_i \in \mathbb{K}^{n-k} \times \mathbb{K}^k$ y Beto recibiese la palabra $\mathbf{c}' = \mathbf{c}'_r\mathbf{c}'_i \in \mathbb{K}^{n-k} \times \mathbb{K}^k$ entonces Beto calcula el síndrome $s'(X) = c'(X) \bmod g(X)$. Si acaso $s'(X) = 0$ entonces Beto acepta a $\mathbf{c}'$ como $\mathbf{c}$ y no reconoce que hubiera habido errores. Si, en cambio, $s'(X) \neq 0$ entonces reconoce que hubo errores y ha de proceder a corregirlos. Se ha de tener una tabla estándar de errores, de distancia de Hamming mínima, para algunos síndromes de manera que cuando se tenga un síndrome particular se lo localice en esa tabla para identificar el error del que proviene, y si no apareciese entonces se procede a rotar la palabra recibida y a multiplicar su síndrome por $X$, reducirlo módulo $g(X)$, y volver a realizar la prueba.

## 8.3 Códigos de Golay

**Definición 8.3** *Sea $\mathbb{K}$ un campo finito. Un código-$(n,k)$ $C$ se dice* perfecto para $t$ errores *si para cualquier palabra de longitud $n$ existe una única palabra de código que diste de ella en a lo sumo $t$. Es decir,*

$$\forall \mathbf{w} \in \mathbb{K}^n\ \exists \mathbf{v} \in C:\ d_n(\mathbf{v},\mathbf{w}) \leq t\ \&\ \forall \mathbf{u} \in C\,(d_n(\mathbf{u},\mathbf{w}) \leq t\ \Rightarrow\ \mathbf{u} = \mathbf{v})\,.$$

Por tanto todo código-$(n,k)$ $C$ perfecto para $t$ errores ha de corregir $t$ errores y en consecuencia su peso mínimo ha de ser $2t+1$.

**Observación 8.3** *Si existe un código-$(n,k)$ $C < \mathbb{F}_2^n$ perfecto para $t$ errores entonces necesariamente*

$$2^{n-k} = \sum_{j=0}^{t} \binom{n}{j}. \tag{25}$$

En efecto, $2^{n-k}$ es el número de clases laterales módulo $C$ en tanto que $\sum_{j=0}^{t}\binom{n}{j}$ es el número de representantes principales de clases con pesos a lo sumo $t$. □

A manera de recíproco, se tiene:

**Observación 8.4** *Si $C < \mathbb{F}_2^n$ es un código-$(n,k)$ que corrige $t$ errores y $2^{n-k} = \sum_{j=0}^{t}\binom{n}{j}$ entonces es perfecto.*

En la tabla 4 presentamos una lista de tercetas de enteros $(n,t,k)$ tales que se cumple (25). En la tercera columna, tercer rengón de ella aparece la terceta $(n,t,k) = (23,3,12)$ por lo cual ha de existir un código-$(23,12)$ $C$ perfecto para $t$ errores.

**Proposición 8.1 (Golay, 1949)** *El código cíclico de longitud 23 con polinomio generador*

$$g(X) = 1 + X^2 + X^4 + X^5 + X^6 + X^{10} + X^{11}$$

*y por consiguiente con polinomio revisor de paridad*

$$h(X) = \frac{X^{23}+1}{g(X)} = 1 + X^2 + X^5 + X^8 + X^9 + X^{10} + X^{11} + X^{12}$$

*es perfecto para 3 errores. Se le llama* código de Golay.

| $(n$ | $t$ | $k)$ | $(n$ | $t$ | $k)$ | $(n$ | $t$ | $k)$ | $(n$ | $t$ | $k)$ | $(n$ | $t$ | $k)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | 1 | 5 | 2 | 1 | 7 | 1 | 4 | 7 | 3 | 1 | 9 | 4 | 1 |
| 11 | 5 | 1 | 13 | 6 | 1 | 15 | 1 | 11 | 15 | 7 | 1 | 17 | 8 | 1 |
| 19 | 9 | 1 | 21 | 10 | 1 | 23 | 3 | 12 | 23 | 11 | 1 | 25 | 12 | 1 |
| 27 | 13 | 1 | 29 | 14 | 1 | 31 | 1 | 26 | 31 | 15 | 1 | 33 | 16 | 1 |
| 35 | 17 | 1 | 37 | 18 | 1 | 39 | 19 | 1 | 41 | 20 | 1 | 43 | 21 | 1 |
| 45 | 22 | 1 | 47 | 23 | 1 | 49 | 24 | 1 | 51 | 25 | 1 | 53 | 26 | 1 |
| 55 | 27 | 1 | 57 | 28 | 1 | 59 | 29 | 1 | 61 | 30 | 1 | 63 | 1 | 57 |
| 63 | 31 | 1 | 65 | 32 | 1 | 67 | 33 | 1 | 69 | 34 | 1 | 71 | 35 | 1 |
| 73 | 36 | 1 | 75 | 37 | 1 | 77 | 38 | 1 | 79 | 39 | 1 | 81 | 40 | 1 |
| 83 | 41 | 1 | 85 | 42 | 1 | 87 | 43 | 1 | 89 | 44 | 1 | 90 | 2 | 78 |
| 91 | 45 | 1 | 93 | 46 | 1 | 95 | 47 | 1 | 97 | 48 | 1 | 99 | 49 | 1 |

Recuadro 4: Tercetas de enteros $(n, t, k)$ tales que $2^{n-k} = \sum_{j=0}^{t} \binom{n}{j}$, con $n \leq 100$.

## 8.4 Códigos de ráfagas

Sea $m \in \mathbb{N}$ un entero positivo. En $\mathbb{F}_2^m$ consideremos la relación de equivalencia:

$$\delta_0 \cdots \delta_m = \boldsymbol{\delta} \sim \boldsymbol{\varepsilon} = \varepsilon_0 \cdots \varepsilon_m \iff \exists j < m \, \forall i < m : \varepsilon_i = \delta_{(i+j) \bmod m}.$$

Es decir, las palabras $\boldsymbol{\delta}$ y $\boldsymbol{\varepsilon}$ son equivalentes si coinciden salvo una rotación de sus símbolos. Una palabra $\boldsymbol{\delta}$ se dice poseer un *patrón de error de ráfaga* de longitud $n$ si para alguna palabra $\boldsymbol{\sigma} \in \mathbb{F}_2^n$ se cumple que $\boldsymbol{\delta} \sim \boldsymbol{\sigma} 0^{(m-n)} = \boldsymbol{\varepsilon}$, es decir, si la parte no nula de la palabra es de longitud a lo sumo $n$ (acaso identificando sus extremos). Si para dos palabras $\boldsymbol{\delta}, \boldsymbol{\varepsilon} \in \mathbb{F}_2^m$ se tiene que el "error" entre ellas $\boldsymbol{\delta} + \boldsymbol{\varepsilon} \in \mathbb{F}_2^m$ es un patrón de error de ráfaga de longitud $n$, entonces se dice que ellas *difieren por una ráfaga de longitud* $n$.

**Proposición 8.2** *Sea $C$ un código-$(n, k)$ cíclico. Supongamos que al transmitir una palabra codificada $\boldsymbol{\delta}$ se han modificado $t$ caracteres.*

1. *Si el síndrome posee un peso de Hamming a lo sumo $t$, entonces ha de coincidir con el patrón de error.*
2. *Si se supone que los errores sólo pueden aparecer en $n-k$ posiciones contiguas, entonces alguna rotación de $\boldsymbol{\delta}$ posee un síndrome con peso de Hamming a lo sumo $t$.*

En efecto, se tiene $e(X) = q(X)\, g(X) + s(X)$, donde $e(X)$ es el polinomio de *error*, $g(X)$ es un generador del código y $s(X)$ es el polinomio de *índrome*. Entonces, $e(X) - s(X) = q(X)\, g(X)$ está en el código. Si el peso de Hamming de $e(X)$ es a lo sumo $t$ entonces $e(X) - s(X)$ posee un peso de Hamming a lo sumo $2t$. Ya que $C$ corrige a lo más $t$ errores, su distancia mínima es $2t + 1$. Por tanto $e(X) - s(X) = 0$ y $e(X) = s(X)$. La segunda aseveración se sigue de ésta inmediatamente. □

## 8.5 Códigos de Reed-Solomon

### 8.5.1 Códigos RS: Como códigos de evaluación

Sea $q$ una potencia de un primo $p$ y $k \in \mathbb{N}$. Sean $a_0, \ldots, a_{n-1} \in \mathbb{F}_q$ $n$ puntos distintos. Se define $\Phi_0 : \mathbb{F}_q^k \to \mathbb{F}_q[X]$ como $\mathbf{c} \mapsto \sum_{j=0}^{k-1} c_j X^j$, es decir como la función que a cada vector de $k$ entradas le asocia el polinomio cuyos coeficientes son esas entradas, y se define

$$\Phi_1 : \mathbb{F}_q^k \to \mathbb{F}_q^n \ , \ \mathbf{c} \mapsto \left( \Phi_0(\mathbf{c})(a_i) = \sum_{j=0}^{k-1} c_j a_i^j \right)_{i=0}^{n-1}.$$

El *código de Reed-Solomon* $\mathrm{RS}_e(q,n,k)$ es $\Phi_1(\mathbb{F}_q^k)$. Una matriz generatriz es pues

$$\left[a_i^j\right]_{(i,j)\in[\![0,n-1]\!]\times[\![0,k-1]\!]} \in \mathbb{F}_q^{n\times k}.$$

$\mathrm{RS}_e(q,n,k)$ es pues un código lineal y su distancia mínima es $n+1-k$ (si $\mathbf{c},\mathbf{d}\in\mathbb{F}_q^k$ son distintos, los polinomios que determinan, $\Phi_0(\mathbf{c})$ y $\Phi_0(\mathbf{d})$, pueden coincidir en a lo sumo $k-1$ puntos, por tanto han de discrepar en al menos $n-k+1$ puntos en $\{a_0,\ldots,a_{n-1}\}$).

**Proposición 8.3** *Dados una potencia de un primo $q$ y $k,n\in\mathbb{N}$ tales que $k\leq n\leq q$, existe un código-$(n,k,n-k+1)$ lineal sobre $\mathbb{F}_q$.*

A saber, basta con dar $\mathrm{RS}_e(q,n,k)$. □

Para $q=2^8=256$, $n=q$ y $k=240$, $\mathrm{RS}_e(256,256,240)$ es un código-$(256,240,17)$ lineal que se utiliza en discos compactos.

Sea $a\in\mathbb{F}_q$ un elemento primitivo, es decir un generador de $\mathbb{F}_q^*$ y sea $n\in\mathbb{N}$ un entero que no sea múltiplo de la característica $p$ de $\mathbb{F}_q$. Entonces $q$ está en el grupo multiplicativo $\mathbb{Z}_n^*$. Sea $m=o_{\mathbb{Z}_n^*}(q)$ el orden de $q$ en $\mathbb{Z}_n^*$. Luego, $q^m=1 \bmod n$. Sea $\rho=a^{\frac{q^m-1}{n}}$. Claramente $\rho^n=1$ en $\mathbb{F}_q$. Así pues la sucesión $\left(\rho^i\right)_{i=0}^{n-1}$ consta de las raíces $n$-ésimas de la unidad en $\mathbb{F}_q$. Al tomar como puntos de evaluación $a_i=\rho^i$, $i\in[\![0,n-1]\!]$, se tendrá que la matriz generatriz de $\mathrm{RS}_e(q,n,k)$ es $\left[\rho^{ij}\right]_{(i,j)\in[\![0,n-1]\!]\times[\![0,k-1]\!]}\in\mathbb{F}_q^{n\times k}$.

### 8.5.2 Códigos RS: Como códigos cíclicos

Si en el campo $\mathbb{F}_q$ consideramos $n=q-1$, entonces las raíces $(q-1)$-ésimas de la unidad en $\mathbb{F}_q$ son todos sus elementos no-nulos, es decir, los elementos del grupo mutiplicativo, y el código de Reed-Solomon será cíclico. Veamos esto con detalle. Seguiremos aquí la presentación hecha en [4].

Sea $q\in\mathbb{N}$ la potencia de un número primo $p$ y sea $\mathbb{K}=\mathbb{F}_q$ el campo de Galois de $q$ elementos, de característica $p$. Sea $a\in\mathbb{F}_q$ un elemento primitivo, vale decir, un generador del grupo cíclico multiplicativo $\mathbb{F}_q^*$. Para $k\leq q-1$, sea $g_{q-1,k}(X)$ el polinomio

$$g_{q-1,k}(X)=\prod_{\kappa=1}^{q-1-k}\left(X-a^\kappa\right).$$

Como $\forall\kappa\in[\![1,q-1-k]\!]$, $(a^\kappa)^{q-1}=\left(a^{q-1}\right)^\kappa=1^\kappa=1$, se tiene $(X-a^\kappa)\,|\,\left(X^{q-1}-1\right)$, y por tanto $g_{q-1,k}(X)|\left(X^{q-1}-1\right)$.

El *código de Reed-Solomon* $\mathrm{RS}(q-1,k)$ de longitud $q-1$ y dimensión $k$ es el código cíclico generado por el polinomio $g_{q-1,k}(X)$. Por tanto:

$$\forall f(X)\in\mathbb{F}_q[X]:[f(X)\in\mathrm{RS}(q-1,k)\iff \mathrm{grd}\,(f)\leq q-2\,\&\,\forall\kappa\in[\![1,q-1-k]\!]:\,f(a^\kappa)=0].$$

Observamos aquí que la última subcondición puede sustituirse por la que impone que $q-1-k$ potencias consecutivas de $a$ son raíces de $f$ (esto porque el código es cíclico). Observamos también que si $a$ no fuese primitivo, entonces ha de generar un subgrupo de $\mathbb{F}_q^*$, por tanto el orden ha de ser un divisor de $q-1$. Se tendría entonces un código de menor tamaño, pero también cíclico.

Al escribir a un elemento $f(X)\in\mathrm{RS}(q-1,k)$ como $f(X)=\sum_{\lambda=0}^{q-2}f_\lambda X^\lambda$, se tiene $\forall\kappa\in[\![1,q-1-k]\!]$

$$0=f(a^\kappa)=\sum_{\lambda=0}^{q-2}f_\lambda a^{\kappa\lambda}=\mathbf{a}_\kappa^T\mathbf{f}$$

donde

$$\mathbf{a}_\kappa=\begin{bmatrix}1\\ a^\kappa\\ \vdots\\ a^{\kappa(q-2)}\end{bmatrix},\quad \mathbf{f}=\begin{bmatrix}f_0\\ f_1\\ \vdots\\ f_{q-2}\end{bmatrix}.$$

Por tanto una matriz revisora de paridad del código $\mathrm{RS}(q-1,k)$ es:

$$\mathrm{RS}^{\perp}_{q-1,k} = \left[a^{\kappa\lambda}\right]_{(\kappa,\lambda)\in[\![1,q-1-k]\!]\times[\![0,q-2]\!]} = \begin{bmatrix} 1 & a & \cdots & a^{(q-2)} \\ 1 & a^2 & \cdots & a^{2(q-2)} \\ \vdots & \vdots & & \vdots \\ 1 & a^{q-1-k} & \cdots & a^{(q-1-k)(q-2)} \end{bmatrix} \in \mathbb{F}_q^{(q-1-k)\times(q-1)}. \qquad (26)$$

**Observación 8.5** *Cualesquiera $q-1-k$ columnas de la matriz $\mathrm{RS}^{\perp}_{q-1,k}$, dada por (26), son linealmente independientes en $\mathbb{F}_q^{q-1-k}$. En otras palabras, al tomar cualesquiera $q-1-k$ columnas de la matriz $\mathrm{RS}^{\perp}_{q-1,k}$, la submatriz resultante es no-singular.*

**Observación 8.6** *El código $\mathrm{RS}(q-1,k)$ tiene distancia mínima $d = q-k$.*

En efecto, por la observación anterior $d \geq q-k$. Por la desigualdad de Singleton $d \leq q-k$. □

**Ejemplo 8.4** *Sean $q = 2^3 = 8$ y $k = 3$. Entonces $q-1-k = 4$. Al tomar como elemento primitivo $a =$ `010`, se tiene que el grupo multiplicativo $\mathbb{F}_8^*$ queda representado como sigue:*

| $\kappa$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $a^\kappa$ | `100` | `010` | `001` | `110` | `011` | `111` | `101` |

*Por otro lado, $g_{73}(X) = (X-a)(X-a^2)(X-a^3)(X-a^4) = a^3 + aX + X^2 + a^3X^3 + X^4$, y el código $\mathrm{RS}(7,3)$ es el generado por el polinomio $g_{73}(X)$.*

Dada una palabra $\sigma \in \mathbb{F}_8^3$, digamos $\sigma =$ `101 001 111`, ésta puede identificarse con el polinomio $\sigma(X) = a^6 + a^2X + a^5X^2$. Al codificarlo, se obtiene el polinomio

$$\sigma(X)\, g_{73}(X) = a^2 + a^4X + a^2X^2 + a^6X^3 + a^6X^4 + a^4X^5 + a^5X^6,$$

el cual polinomio representa a la 7-palabra $\tau =$ `001 011 001 101 101 011 111`. Así pues, el código $\mathrm{RS}(7,3)$ establece la correspondencia $\sigma \mapsto \tau$. □

## 8.6 Decodificación de Reed-Solomon

Sea $q$ una potencia de un número primo y sea $k \leq q-1$. Supongamos que para una palabra $\sigma \in \mathbb{F}_q^k$, su código correspondiente a $\mathrm{RS}(q-1,k)$ es $f(X) = \sum_{\lambda=0}^{q-2} f_\lambda X^\lambda \in \mathbb{F}_q[X]$, pero que, al ser transmitido, el destinatario hubiera recibido $h(X) = \sum_{\lambda=0}^{q-2} h_\lambda X^\lambda$. El propósito de la decodificación es recuperar $f(X)$ a partir de $h(X)$ o, en otras palabras, calcular el error $e(X) = h(X) - f(X)$. Naturalmente, la recuperación de $\sigma$ a partir de $f(X)$ se hace dividiendo a este último entre el polinomio $g_{q-1,k}(X)$.

Para cada $\kappa \in [\![1, q-1-k]\!]$ denotemos por $s_\kappa$ al síndrome $s_\kappa = e(a^\kappa) = h(a^\kappa)$ (pues $f(a^\kappa) = 0$).

Inicialmente, consideremos el caso $k = 1$.

Supongamos que el error entre $h(X)$ y $f(X)$ ocurre solamente en un "byte", es decir, en un solo coeficiente, digamos $h_\lambda$. Entonces los síndromes han de satisfacer $s_1 = e_\lambda a^\lambda$ y $s_2 = e_\lambda a^{2\lambda}$. Por tanto $a^\lambda = \frac{s_2}{s_1}$ y, en consecuencia, $e_\lambda = \frac{s_1^2}{s_2}$. Así pues:

> **Decodificación de Reed-Solomon para $k = 1$.** Si sólo hay un error en un solo coeficiente, la posición en la que ocurre es $\lambda = \log_a\left(\frac{s_2}{s_1}\right)$ y el coeficiente del error ahí es $e_\lambda = \frac{s_1^2}{s_2}$.

Para $k > 1$, sea $E \subset [\![0, q-2]\!]$ el conjunto de índices $\lambda$ tales que $e_\lambda \neq 0$. Entonces se tiene el sistema de ecuaciones lineales:

$$\forall \kappa \in [\![0, q-1-k]\!] : \quad s_\kappa = \sum_{\lambda\in E} e_\lambda a^{\kappa\lambda} \qquad (27)$$

con soluciones únicas siempre que $2\,\mathrm{card}\,(E) \leq q-1-k$. Sin embargo éste no determina al conjunto de índices $E$.

Para determinarlo es necesaria una labor suplementaria. Se define el *polinomio localizador de errores* como

$$\rho(X) = \prod_{\lambda \in E} (X - a^{-\lambda}). \tag{28}$$

el cual es un polinomio mónico. Entonces las raíces de $\rho(X)$ son precisamente las potencias $a^{-\lambda}$ con $\lambda \in E$. También se define el *polinomio evaluador de errores* como

$$\omega(X) = \sum_{\lambda \in E} e_\lambda \prod_{\ell \in E - \{\lambda\}} (X - a^{-\ell}). \tag{29}$$

Se tiene $\mathrm{grd}\,(\rho(X)) = \mathrm{card}\,(E)$ y $\mathrm{grd}\,(\omega(X)) = \mathrm{card}\,(E) - 1$. Se supondrá $2\,\mathrm{card}\,(E) \leq q-1-k$.

**Proposición 8.4** *Los polinomios $\rho(X)$ y $\omega(X)$ son primos relativos y*

$$\forall \lambda \in E:\ e_\lambda = \frac{\omega(a^{-\lambda})}{\rho'(a^{-\lambda})}, \tag{30}$$

*donde $\rho'(X)$ es la derivada formal del polinomio $\rho(X)$.*

En efecto, para lo primero observamos que $\rho(X)$ y $\omega(X)$ no pueden tener raíces comunes:

$$\begin{aligned} \rho(x) = 0 &\implies x = a^{-\lambda} \ \text{ para alguna } \lambda \in E \\ &\implies \omega(a^{-\lambda}) = e_\lambda \prod_{\ell \in E - \{\lambda\}} (a^{-\lambda} - a^{-\ell}) \\ &\implies \omega(a^{-\lambda}) \neq 0. \end{aligned}$$

Para lo segundo, por la fórmula de Leibniz, la derivada formal de $\rho(X)$ es

$$\rho'(X) = \sum_{\lambda \in E} \prod_{\ell \in E - \{\lambda\}} (X - a^{-\ell})$$

y por tanto $\rho'(a^{-\lambda}) = \prod_{\ell \in E - \{\lambda\}} (a^{-\lambda} - a^{-\ell})$. Así resulta la expresión (30). □

Si se determinara el polinomio localizador de errores $\rho(X)$ entonces sus raíces determinan a su vez el conjunto $E$. Consideremos un tercer polinomio, llamado *de síndromes*, pues éstos son sus coeficientes:

$$\sigma(X) = s_1 + s_2 X + \cdots + s_{q-1-k} X^{q-2-k} = \sum_{\kappa=0}^{q-2-k} s_{\kappa+1} X^\kappa.$$

**Proposición 8.5** *Existe un polinomio $\mu(X) \in \mathbb{F}_q[X]$ tal que se cumple la* ecuación clave*:*

$$\rho(X)\sigma(X) = \mu(X)\, X^{q-1-k} - \omega(X). \tag{31}$$

*O equivalentemente:*

$$\rho(X)\sigma(X) = -\omega(X) \bmod X^{q-1-k}.$$

En efecto, un cálculo directo da:

$$\begin{aligned} \sigma(X) &= \sum_{\kappa=0}^{q-2-k} e(a^{\kappa+1})\, X^\kappa \\ &= \sum_{\kappa=0}^{q-2-k} \left[ \sum_{\lambda \in E} e_\lambda a^{(\kappa+1)\lambda} \right] X^\kappa \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\lambda\in E} e_\lambda a^\lambda \left[\sum_{\kappa=0}^{q-2-k} \left(a^\lambda X\right)^\kappa\right] \\
&= \sum_{\lambda\in E} e_\lambda \left[\frac{\left(a^\lambda X\right)^{q-1-k}-1}{X-a^{-\lambda}}\right], \qquad (32)
\end{aligned}
$$

de donde

$$
\begin{aligned}
\rho(X)\sigma(X) &= \sum_{\lambda\in E} e_\lambda \left[\left(a^\lambda X\right)^{q-1-k}-1\right] \prod_{\ell\in E-\{\lambda\}} (X-a^{-\ell}) \\
&= \sum_{\lambda\in E} e_\lambda \left(a^\lambda X\right)^{q-1-k} \prod_{\ell\in E-\{\lambda\}} (X-a^{-\ell}) - \sum_{\lambda\in E} e_\lambda \prod_{\ell\in E-\{\lambda\}} (X-a^{-\ell}) \\
&= X^{q-1-k}\left[\sum_{\lambda\in E} e_\lambda a^{(q-1-k)\lambda} \prod_{\ell\in E-\{\lambda\}} (X-a^{-\ell})\right] - \omega(X) \\
&= X^{q-1-k}\,\mu(X) - \omega(X).
\end{aligned}
$$

□

Así pues, para decodificar, habiendo calculado el polinomio de síndromes $\sigma(X)$, se ha de determinar a los polinomios localizador y evaluador de errores $\rho(X)$ y $\omega(X)$ de manera que se cumpla la ecuación clave (31).

### 8.6.1 Método PGZ

Veamos una primera forma de resolver la ecuación clave, mediante el llamado *decodificador de Peterson-Gorenstein-Zierler* (PGZ).

Supongamos que hubieran ocurrido $m$ errores, con $2m \le q-1-k$. Escribamos a los polinomios localizador y evaluador de errores como:

$$
\begin{aligned}
\rho(X) &= r_0 + r_1 X + \cdots + r_{m-1}X^{m-1} + X^m \\
\omega(X) &= u_0 + u_1 X + \cdots + u_{m-1}X^{m-1}
\end{aligned}
$$

El producto $\rho(X)\sigma(X)$ es de grado a lo sumo $m+q-2-k$, pero para $j \in [\![m, q-2-k]\!]$ la ecuación clave implica que el coeficiente correspondiente en ese producto es nulo. Por la manera en la que se multiplica a los polinomios, se ha de tener pues:

$$
j \in [\![m, q-2-k]\!] \;\Rightarrow\; \sum_{\ell=0}^{m} r_\ell s_{j+1-\ell} = 0.
$$

Puesto que $r_m = 1$, se tiene:

$$
j \in [\![m, q-2-k]\!] \;\Rightarrow\; \sum_{\ell=0}^{m-1} r_\ell s_{j+1-\ell} = -s_{j+1-m}.
$$

Estas condiciones plantean, de forma matricial, el sistema de ecuaciones:

$$
\begin{bmatrix}
s_{m+1} & s_m & s_{m-1} & \cdots & s_2 \\
s_{m+2} & s_{m+1} & s_m & \cdots & s_3 \\
s_{m+3} & s_{m+2} & s_{m+1} & \cdots & s_4 \\
\vdots & \vdots & \vdots & & \vdots \\
s_{q-k-1} & s_{q-k-2} & s_{q-k-3} & \cdots & s_{q-k-(1+m)}
\end{bmatrix}
\begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{m-1} \end{bmatrix}
= -
\begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_{q-k-(1+m)} \end{bmatrix} \qquad (33)
$$

o sea $\mathbf{S}\cdot\mathbf{r}=\mathbf{s}$, donde $\mathbf{S}\in\mathbb{F}_q^{(q-k-(1+m))\times m}$, $\mathbf{s}\in\mathbb{F}_q^{q-k-(1+m)}$ y $\mathbf{r}\in\mathbb{F}_q^m$ es el vector de coeficientes de $\rho(X)$ y hace aquí el papel de incógnita. El sistema (33) está sobredimensionado (hay más condiciones que incógnitas), puede pues resolverse tomando solamente sus primeros $m$ renglones. Queda el sistema:

$$\begin{bmatrix} s_{m+1} & s_m & s_{m-1} & \cdots & s_2 \\ s_{m+2} & s_{m+1} & s_m & \cdots & s_3 \\ s_{m+3} & s_{m+2} & s_{m+1} & \cdots & s_4 \\ \vdots & \vdots & \vdots & & \vdots \\ s_{2m} & s_{2m-1} & s_{2m-2} & \cdots & s_{m+1} \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{m-1} \end{bmatrix} = - \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \\ s_m \end{bmatrix} \tag{34}$$

$$\mathbf{S}_m\cdot\mathbf{r} = -\mathbf{s}_m$$

Como los síndromes son no nulos, y $\mathbf{S}_m$ es una matriz definida por diagonales, se tiene que es no-singular. Por tanto el polinomio localizador de errores se obtiene como $\mathbf{r}=-\mathbf{S}_m^{-1}\mathbf{s}_m$.

Habiendo así localizado el conjunto $E$, el sistema (27) permite entonces calcular el error $e(X)$, con lo cual se ha de completar el proceso de decodificación.

También puede verse que si $m'\in[\![m+1,q-1-k]\!]$ entonces $\mathbf{S}_{m'}$ es singular. Así pues, si se desconociera cuál es el valor de $m$, entonces éste se obtendría como el máximo $m'$ tal que $\mathbf{S}_{m'}$ es no-singular.

### 8.6.2 Método de Euclides

Escribiendo la ecuación clave como

$$\mu(X)\,X^{q-1-k}+(-\rho(X))\sigma(X)=-\omega(X)$$

tenemos que $\omega(X)$ es un máximo común divisor (salvo por el signo) de $X^{q-1-k}$ y el polinomio de síndromes $\sigma(X)$. Así pues, utilizando el Algoritmo de Euclides, al calcular una forma extendida del Máximo Común Divisor y expresar a $-\omega(X)$ como una combinación lineal de $X^{q-1-k}$ y $\sigma(X)$ se obtendrá los polinomios $\mu(X)$ y el localizador de errores $\rho(X)$. Este último determina el conjunto $E$ y el sistema (27) permite entonces calcular el error $e(X)$, con lo cual se ha de completar el proceso de decodificación.

## 8.7 Códigos BCH

Los códigos de *Bose, Chaudhuri y Hocquenghem* (BCH), se construyen a partir de los de Reed-Solomon pero se les restringe a estar en subcampos del campo $\mathbb{F}_q$ original.

Recordamos que un *subcampo* $\mathbb{K}_0$ de un campo $\mathbb{K}_1$ es un subconjunto $\mathbb{K}_0\subset\mathbb{K}_1$ tal que $0,1\in\mathbb{K}_0$ y $\mathbb{K}_0$ es cerrado bajo las operaciones de adición, multiplicación e inversos aditivos y multiplicativos. En tal caso, con las operaciones de $\mathbb{K}_1$, $\mathbb{K}_0$ es un campo.

Si $x\in\mathbb{K}_1$, el polinomio $p_x(X)\in\mathbb{K}_0[X]$, con coeficientes en el subcampo $\mathbb{K}_0$, de grado mínimo tal que $p_x(x)=0$ se dice ser el *polinomio mínimo* de $x$ respecto a $\mathbb{K}_0$. Se tiene que $p_x(X)$ es irreducible sobre $\mathbb{K}_0$ y divide a cualquier polinomio en $\mathbb{K}_0[X]$ que tenga a $x$ como raíz.

Un campo finito $\mathbb{F}_{q_1}$ sólo es tal si $q_1=p^{n_1}$ es una potencia de un primo $p$, el cual es la característica de $\mathbb{F}_{q_1}$: $px=0$ para cada $x\in\mathbb{F}_{q_1}$. Ahora, si $\mathbb{F}_{q_0}$ es un subcampo de $\mathbb{F}_{q_1}$ entonces $q_0=p^{n_0}$ con $n_0\le n_1$, pero aún más: el grupo multiplicativo $\mathbb{F}_{q_0}^*$ debe ser un subgrupo de $\mathbb{F}_{q_1}^*$, por tanto $(p^{n_0}-1)|(p^{n_1}-1)$ lo cual implica $n_0|n_1$. En resumen: $\mathbb{F}_{q_0}$ es un subcampo de $\mathbb{F}_{q_1}$ si y sólo si $q_0=p^{n_0}$, $q_1=p^{n_1}$ y $n_0|n_1$.

### 8.7.1 Una primera presentación

Como en la sección anterior, aquí seguiremos la presentación hecha en [4].

Sea $q_1$ una potencia de un primo $p$, sea $k\le q_1-1$ y $q_0$ tal que $\mathbb{F}_{q_0}$ sea un subcampo de $\mathbb{F}_{q_1}$. El correspondiente *código BCH* es el $\mathbb{F}_{q_0}$-subespacio lineal $C$ de $\mathrm{RS}(q_1-1,k)$ tal que $C\subset\mathbb{F}_{q_0}[X]$.

Sea $g_{q_1-1,k}(X)=\prod_{i=1}^{q_1-1-k}(X-a^i)$ el polinomio generador de $\mathrm{RS}(q_1-1,k)$, donde $a$ es un elemento primitivo de $\mathbb{F}_{q_1}$. Entonces $g_{q_1-1,k}(a^i)=0$ siempre que $i\in[\![1,q_1-1-k]\!]$. Por tanto, si un polinomio $\pi(X)\in\mathbb{F}_{q_1}[X]$ está en el código BCH $C$ entonces $\pi(X)\in\mathbb{F}_{q_0}[X]$ y $\pi(a^i)=0$ para todo $i\in[\![1,q_1-1-k]\!]$.

De su definición, no se desprende inmediatamente cuáles serán la longitud, la dimensión y la distancia mínima de un código BCH $C$. Sin embargo, la distancia mínima $q_1-1-k$ de $\mathrm{RS}(q_1-1,k)$ se dice ser la *distancia prevista* del código $C$.

**Observación 8.7** *Sean $q = p^n$, $x \in \mathbb{F}_q$ y $p_x(X) \in \mathbb{F}_p[X]$ el polinomio mínimo respecto a $\mathbb{F}_p$.*

- *Como $x^q = 1$ se tiene $p_x(X) | (X^q - 1)$ en $\mathbb{F}_p[X]$.*
- *Por otro lado, sea $n_x = \operatorname{grd} p_x(X)$ el grado del polinomio mínimo. Se tiene que el conjunto $L = \{1, x, \ldots, x^{n_x - 1}\}$ es linealmente independiente, respecto a $\mathbb{F}_p$, pero $L \cup \{x^{n_x}\}$ no lo es. Por tanto $n_x \leq n$.*
- *En particular, si $x$ es primitivo, entonces $n_x = n$.*

**Observación 8.8** *Sean $n_0, n_1 \in \mathbb{N}$ tales que $n_0 | n_1$, $p$ un número primo, $q_0 = p^{n_0}$ y $q_1 = p^{n_1}$. Sea $a_1 \in \mathbb{F}_{q_1}$ un elemento primitivo en ese campo. Sea $m = \frac{q_1 - 1}{q_0 - 1}$ y $a_0 = a_1^m$. Entonces $a_0$ es generador de un subgrupo $\langle a_0 \rangle$, isomorfo a $\mathbb{F}_{q_0}^*$, de $\mathbb{F}_{q_1}^*$. De hecho, $\langle a_0 \rangle \cup \{0\}$ es un subcampo de $\mathbb{F}_{q_1}$, isomorfo a $\mathbb{F}_{q_0}$.*

*En estas condiciones, si $p_{a_0}(X)$ es el polinomio mínimo de $a_0$ respecto a $\langle a_0 \rangle \cup \{0\}$, entonces $\mathbb{F}_{q_1}/(p_{a_0}(X))$ es isomorfo a $\mathbb{F}_{q_0}$. En consecuencia, $\operatorname{grd} p_{a_0}(X) = n_0$.*

Supongamos ahora que $\mathbb{F}_{q_0}$ es un subcampo de $\mathbb{F}_{q_1}$. Veamos cómo calcular polinomios mínimos de elementos en $\mathbb{F}_{q_1}$ respecto a $\mathbb{F}_{q_0}$.

Se dice que dos elementos $y, z \in \mathbb{F}_{q_1}$ son *conjugados* respecto a $\mathbb{F}_{q_0}$, si existen $x \in \mathbb{F}_{q_1}$ e $i, j \in [\![0, \frac{n_1}{n_0} - 1]\!]$ tales que $y = x^{q_0^i}$ y $z = x^{q_0^j}$. Esto introduce una relación de equivalencia en $\mathbb{F}_{q_1}$.

Para cada $x \in \mathbb{F}_{q_1}$, sea $S_{q_0}(x) = \left\{ x^{q_0^j} \right\}_{j=0}^{\frac{n_1}{n_0} - 1}$ (naturalmente para $j = \frac{n_1}{n_0}$, $q_0^j = (p^{n_0})^j = p^{n_1} = q_1$, por tanto $x^{q_0^j} = x^{q_1} = x$). Así cada $S_{q_0}(x)$ es una clase de conjugación y posee $\frac{n_1}{n_0}$ elementos.

Debido a que la transformación $z \mapsto z^{q_0}$ es un homomorfismo $\mathbb{F}_{q_1} \to \mathbb{F}_{q_1}$ se tiene:

**Observación 8.9** *Si $\pi(X) \in \mathbb{F}_{q_0}[X]$ entonces rige la implicación siguiente:*

$$\forall x \in \mathbb{F}_{q_1} : \ [\pi(x) = 0 \implies \pi(x^{q_0}) = 0] .$$

Por tanto, los conjugados de raíces de polinomios con coeficientes en el subcampo son también raíces de esos polinomios. Así, si $a$ es primitivo en $\mathbb{F}_{q_1}$ y $\pi(X)$ está en el código BCH entonces las potencias $a^i$, con $i \in [\![1, q_1 - 1 - k]\!]$, y sus conjugados son raíces de $\pi(X)$.

**Proposición 8.6** *Para cada $x \in \mathbb{F}_{q_1}$, el polinomio mínimo $p_x(X)$ respecto a $\mathbb{F}_{q_0}$ queda caracterizado como*

$$p_x(X) = \prod_{y \in S_{q_0}(x)} (X - y) . \tag{35}$$

Se demuestra la proposición viendo que al expandir la expresión a la derecha de (35) resulta un polinomio con coeficientes en $\mathbb{F}_{q_0}$, irreducible ahí y que es el de grado mínimo que se anula en $x$. □

Sean $a$ un elemento primitivo en $\mathbb{F}_{q_1}$ y $g_{q_1 - 1, k}(X) = \prod_{i=1}^{q_1 - 1 - k} (X - a^i)$ el generador de $\mathrm{RS}(q_1 - 1, k)$. Si $\pi(X)$ está en el correspondiente código BCH, entonces $\pi(a^i) = 0$, con $i \in [\![1, q_1 - 1 - k]\!]$. Por tanto, cada uno de los polinomios mínimos $p_{a^i}(X)$, respecto al subcampo $\mathbb{F}_{q_0}$, divide a $\pi(X)$. En consecuencia, el mínimo común múltiplo $\operatorname{mcm} \{p_{a^i}(X)\}_{j=1}^{q_1 - 1 - k}$ también divide a $\pi(X)$. En particular, debe dividir también al generador $g_{q_1 - 1, k}(X)$. Por la minimalidad de este último, necesariamente se ha de tener: $g_{q_1 - 1, k}(X) = \operatorname{mcm} \{p_{a^i}(X)\}_{j=1}^{q_1 - 1 - k}$.

### 8.7.2 Una segunda presentación

Sea $q$ una potencia de un primo $p$, $n \in \mathbb{N}$ un entero que no es múltiplo de $p$, $k \leq n - 2$ y $\rho$ una raíz $n$-ésima de la unidad en $\mathbb{F}_q$. Sean $R = \left(\rho^j\right)_{j=\ell}^{\ell + n - k - 1}$ una colección de $n - k$ raíces $n$-ésimas consecutivas de la unidad y $g_{\ell, n, k}(X) = \operatorname{mcm} \left( \left( p_{\rho^j}(X) \right)_{\rho^j \in R} \right)$ el mínimo común múltiplo de los polinomios mínimos, respecto a $\mathbb{F}_q$, de los elementos en $R$. El código cíclico generado por $g_{\ell, n, k}(X)$ se dice ser el *código BCH* generado por $R$.

Los códigos BCH de longitud $n = q - 1$ son precisamente los códigos de Reed-Solomon (en este caso, los polinomios mínimos son $(X - \rho^i)$, $i \leq n$).

**Ejemplo 8.5** *Constrúyase un código BCH que corrija errores dobles.*

Como se quiere corregir hasta 2 errores la distancia mínima debe ser 5. Se opta por un código de Reed-Solomon. Por tanto $q-k=5$. Tómese $q=7$ y $k=2$. Entonces $n=q-1=6$ y $n-k=4$. En $\mathbb{F}_7$ un elemento primitivo es $\rho=3$. Así, sea $R=\left(\rho^j\right)_{j=1}^{1+n-k-1}=\{3,3^2,3^3,3^4\}=\{3,2,6,4\}$. El generador es pues

$$g_{1,6,2}(X)=(X-3)(X-2)(X-6)(X-4)=4+2X+3X^2+6X^3+X^4.$$

La matriz generatriz es

$$G=\begin{bmatrix}4&2&3&6&1&0\\0&4&2&3&6&1\end{bmatrix}$$

y el polinomio revisor de paridad es

$$h(X)=\frac{X^6-1}{g_{1,6,2}(X)}=5+X+X^2$$

lo que define la matriz de paridad

$$H=\begin{bmatrix}1&1&5&0&0&0\\0&1&1&5&0&0\\0&0&1&1&5&0\\0&0&0&1&1&5\end{bmatrix}$$

## 8.8 Terceras listas

### 8.8.1 Ejercicios

1. Sea $a$ un elemento primitivo en un campo $\mathbb{F}_q$. Demuestre las siguientes dos implicaciones:

$$\kappa\neq 0 \bmod (q-1) \implies \sum_{\lambda=0}^{q-2} a^{\kappa\lambda}=0$$

$$\kappa= 0 \bmod (q-1) \implies \sum_{\lambda=0}^{q-2} a^{\kappa\lambda}=q-1$$

2. Sea $g(X)=1+X+X^2+X^4+X^5+X^8+X^{10}$ un polinomio generador de un código cíclico-$[15,5]$. Calcule el polinomio de paridad $h(X)$ de $g(X)$.

3. Sea $g(X)\in\mathbb{F}_2[X]$ un polinomio de grado $m\in\mathbb{N}$ que no sea un binomio, y sea $n$ el entero mínimo tal que $g(X)$ divide a $X^n+1$. Muestre que el código cíclico de longitud $n$ generado por $g(X)$ tiene peso mínimo al menos 3.

4. Sea $C$ un código cíclico-$[n,k]$ binario. Muestre que si $n$ es impar y $X+1$ no divide a $g(X)$ entonces el vector constante 1, $1^{(n)}$, está en el código $C$.

5. Sea $g(X)=1+X^4+X^6+X^7+X^8$ un polinomio generador de un código cíclico-$[15,7]$. Calcule el síndrome del polinomio $a(X)=1+X+X^5+X^{14}$ y decida si $a(X)$ está o no en el código.

6. Sea $C$ el código-$[15,11]$ cíclico binario generado por el polinomio $g(X)=1+X+X^4$, y suponga que se ha recibido el polinomio $c(X)=1+X+X^4+X^{10}$ cuando se ha utilizado una codificación no-sistemática. Recupere el mensaje $a(X)$ que le dió origen.

7. Sea $H_m\in\mathbb{F}_2^{(2^m-1)\times m}$ la matriz revisora de paridad del código de Hamming-$(2^m-1,2^m-m-1)$, dado en la definición 4.7, y sea $C_m\subset\mathbb{F}_2^{2^m-1}$ el código cuya generatriz es $H_m$.

   A. Decida si acaso $C_m$ es un código cíclico.

   B. Muestre que para cada $\mathbf{u}\in C_m-\{\mathbf{0}\}$ y para cada $k\in[\![1,2^m-2]\!]$, $\rho_{2^m-1}^k(\mathbf{u})\neq\mathbf{u}$, donde $\rho_{2^m-1}$ es la rotación a la izquierda de longitud $2^m-1$.

8. Sea $H_m \in \mathbb{F}_2^{(2^m-1)\times m}$ la matriz revisora de paridad del código de Hamming-$(2^m-1, 2^m-m-1)$, dado en la definición 4.7, y sea $C_m \subset \mathbb{F}_2^{2^m-1}$ el código cuya generatriz es $H_m$.

A. Muestre que las palabras de código no-nulas de $C_m$ poseen todas un peso de Hamming $2^{m-1}$.

B. Sea $M \in \mathbb{F}_2^{(2^m-1)\times(2^m-1)}$ la matriz tal que la entrada $m_{ij}$ es la $j$-ésima entrada del $i$-ésimo vector no nulo en $C_m$. De acuerdo con el inciso anterior en cada renglón hay $2^{m-1}$ 1's. Muestre que también en cada columna hay $2^{m-1}$ 1's.

9. *Código* `CRC-16`. Sea $g(X) = 1 + X^2 + X^{15} + X^{16} \in \mathbb{F}_2[X]$, el llamado polinomio `CRC-16`. Se tiene que, en $\mathbb{F}_2[X]$:

$$g(X) = (X+1)\, h(X) = (X+1)\,(1 + X + X^{15})$$

y $h(X)$ es primitivo.

A. ¿Cuál es el mínimo $n \in \mathbb{N}$ tal que $g(X)$ genera un código cíclico de longitud $n$?

B. Diseñe un algoritmo, a nivel de bits, para realizar la codificación no-sistemática en ese código.

10. *Código* `CRC-CCITT`. Sea $g(X) = 1 + X^5 + X^{12} + X^{16} \in \mathbb{F}_2[X]$, el llamado polinomio `CRC-CCITT`. Se tiene que, en $\mathbb{F}_2[X]$:

$$g(X) = (X+1)\, h(X) = (X+1)\,(1 + X + X^2 + X^3 + X^4 + X^{12} + X^{13} + X^{14} + X^{15})$$

y $h(X)$ es primitivo.

A. ¿Cuál es el mínimo $n \in \mathbb{N}$ tal que $g(X)$ genera un código cíclico de longitud $n$?

B. Diseñe un algoritmo, a nivel de bits, para realizar la codificación no-sistemática en ese código.

11. El código de Golay es un código-$(23, 12)$ perfecto de distancia mínima 7.

A. Muestre que hay palabras de código, es decir, vectores en el código de Golay, de peso 16.

B. Muestre que hay un mismo número de vectores en el código de peso 7 que de peso 16.

12. Encuentre el polinomio generador de un código de BCH de longitud 31 que corrija errores dobles y el de otro que corrija errores triples.

13. Encuentre el polinomio generador de un código de Reed-Solomon con símbolos en $\mathbb{F}_{2^5}$ que corrija errores dobles.

14. ¿Cuántos códigos cíclicos de longitud 8 hay sobre $\mathbb{F}_3$?

### 8.8.2 Programas

1. Escriba un programa que reciba un polinomio $g(X) \in \mathbb{F}_2[X]$ de grado $m \in \mathbb{N}$, y encuentre un entero $n > m$ tal que $g(X)|(X^n + 1)$ en $\mathbb{F}_2[X]$, para que posteriormente realice las funciones siguientes:

A. Calcule la matriz generatriz del código-$[n, n-m]$ cíclico $C$ generado por $g(X)$.

B. Calcule la matriz revisora de paridad del código $C$.

C. Calcule a todos los polinomios en el código $C$.

2. Escriba un programa que reciba un polinomio $g(X) \in \mathbb{F}_2[X]$ de grado $m \in \mathbb{N}$, localice primeramente un entero $n > m$ tal que $g(X)|(X^n + 1)$, y luego realice las funciones siguientes:

A. Construya un arreglo estándar del código-$[n, n-m]$ cíclico $C$ generado por $g(X)$. Para esto a cada polinomio $e(X) = \sum_{\nu=0}^{n} \varepsilon_\nu X^\nu \in \mathbb{F}_2[X]$ le debe asociar su síndrome $e(X) \bmod g(X)$.

B. *Codificación no-sistemática*: Dado $a(X) \in \mathbb{F}_2[X]$ de grado a lo sumo $n-m-1$ produce su código $c(X) = a(X)g(X)$.

C. *Decodificación no-sistemática*: Dado $c(X) \in \mathbb{F}_2[X]$ de grado $n-1$, decide si está en el código en cuyo caso recupera $a(X)$ tal que $c(X) = a(X)g(X)$, mas si no estuviera, bajo la suposición de que hay menos de $d$ errores, donde $2d+1$ es la distancia mínima de $C$, los corrige y recupera el mensaje original $a(X)$.

D. *Codificación sistemática*: Dado $a(X) \in \mathbb{F}_2[X]$ de grado a lo sumo $n-m-1$ produce su código $c(X) = -[X^m a(X), \bmod g(X)] + X^m a(X)$.

E. *Decodificación sistemática*: Dado $c(X) \in \mathbb{F}_2[X]$ de grado $n-1$, decide si está en el código en cuyo caso recupera $a(X)$ tal que $c(X) = -[X^m a(X) \bmod g(X)] + X^m a(X)$, mas si no estuviera, bajo la suposición de que hay menos de $d$ errores, los corrige y recupera el mensaje original $a(X)$.

3. *Estructura multiplicativa de $\mathbb{F}_{2^k}$.* Escriba un programa que reciba un primo $p$, un entero positivo $n$ y calcule las tablas de adición y de producto de $\mathbb{F}_{p^n}$, donde sus elementos se ven como palabras de longitud $n$ de símbolos en $\mathbb{F}_p$.

4. *Codificación de Reed-Solomon.* Escriba un programa que, dados los parámetros $q$ (potencia de un primo) y $k \in \mathbb{N}$, reciba como entrada una palabra $\sigma \in \mathbb{F}_q^k$ y calcule su código $\tau \in \mathbb{F}_q^{q-1}$ de acuerdo con el código $\mathrm{RS}(q-1,k)$. (Vea el ejemplo 8.4.)

5. *Decodificación PGZ de Reed-Solomon.* Escriba un programa que, dados los parámetros $q$ (potencia de un primo) y $k \in \mathbb{N}$, reciba como entrada una palabra $h \in \mathbb{F}_q^{q-1}$, calcule el elemento $f \in \mathrm{RS}(q-1,k)$ más cercano a ella, de acuerdo con el procedimiento descrito en la sección 8.6.1, y luego recupere la palabra $\sigma \in \mathbb{F}_q^k$ codificada.

6. *Decodificación mediante Euclides de Reed-Solomon.* Escriba un programa que, dados los parámetros $q$ (potencia de un primo) y $k \in \mathbb{N}$, reciba como entrada una palabra $h \in \mathbb{F}_q^{q-1}$, calcule el elemento $f \in \mathrm{RS}(q-1,k)$ más cercano a ella, de acuerdo con el procedimiento descrito en la sección 8.6.2, y luego recupere la palabra $\sigma \in \mathbb{F}_q^k$ codificada.

# Referencias


[1] Jiri Adamek. *Foundations of Coding: Theory and Applications of Error-Correcting Codes, with an Introduction to Cryptography and Information.* John Wiley & Sons, Inc., New York, NY, USA, 1991.

[2] E. F. Assmus, Jr. and J. D. Key. *Designs and their Codes.* Cambridge: Cambridge University Press, 1992. Cambridge Tracts in Mathematics, Vol. 103 (Second printing with corrections, 1993).

[3] E. F. Assmus, Jr. and J. D. Key. Designs and codes: an update. *Des. Codes Cryptogr.*, 9:7–27, 1996.

[4] Mario Blaum. *A course on error-correcting codes.* Hitachi Global Storage Technologies, San Jose, CA, 2008.

[5] Arnaldo García and Henning Stichtenoth. *Topics in Geometry, Coding Theory and Cryptography (Algebra and Applications).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[6] Paul Garrett. *The Mathematics of Coding Theory.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2003.

[7] D. C. Hankerson, Gary Hoffman, D. A. Leonard, Charles C. Lindner, K. T. Phelps, C. A. Rodger, and J. R. Wall. *Coding Theory and Cryptography: The Essentials.* Marcel Dekker, Inc., New York, NY, USA, 2000.

[8] W. C. Huffman and Richard A. Brualdi. *Handbook of Coding Theory.* Elsevier Science Inc., New York, NY, USA, 1998.

[9] Harald Niederreiter and University Press Singapore. *Coding Theory and Cryptology*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2002.

[10] Ron Roth. *Introduction to Coding Theory*. Cambridge University Press, New York, NY, USA, 2006.

[11] M. R. Schroeder. *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity (Springer Series in Information Sciences)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[12] Wade Trappe and Lawrence C. Washington. *Introduction to Cryptography with Coding Theory (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2005.

[13] Jacobus H. van Lint. *Introduction to Coding Theory*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1997.

[14] Jacobus H. van Lint and Richard M. Wilson. On the minimum distance of cyclic codes. *IEEE Transactions on Information Theory*, 32(1):23–40, 1986.

[15] Lawrence C. Washington and Wade Trappe. *Introduction to Cryptography: With Coding Theory*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2002.