

# Códigos y criptografía

José Carlos García



22 de febrero de 2016



# Índice general

<b>1. Introducción</b>	<b>5</b>
1.1. Cuerpos finitos . . . . .	5
<b>2. Códigos autocorrectores</b>	<b>7</b>
2.1. Parámetros de un código . . . . .	7
2.2. Códigos lineales . . . . .	8
2.2.1. Códigos de Hamming . . . . .	11
2.3. Algunos códigos buenos . . . . .	11
2.4. Códigos cíclicos . . . . .	11
<b>3. Criptografía</b>	<b>13</b>
3.1. Criptosistemas simétricos . . . . .	13
3.2. Criptosistemas de clave pública . . . . .	13



# Capítulo 1

## Introducción

### 1.1. Cuerpos finitos

**Definición 1** (Cuerpo). Un **cuerpo** es un anillo conmutativo con unidad en el que todo elemento distinto de 0 tiene inverso.

**Ejemplo 1.**  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  es anillo conmutativo con unidad, además, si  $n$  es primo entonces  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  es un cuerpo.

La suma y el producto de  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  se define como el resto de la suma de los elementos, para el producto de forma análoga.

**Teorema 1.** Si un  $K$  es finito, entonces  $\text{card}(K) = p^r$  con  $p \in \mathbb{P}$  y  $r \in \mathbb{N}$

**Teorema 2.** Dado  $p \in \mathbb{P}, r \in \mathbb{N}$  entonces existe un cuerpo  $K$  tal que  $\text{card}(K) = p^r$ . Además, dos cuerpos finitos con el mismo cardinal son isomorfos.

**Definición 2** (Cuerpo finito de  $q$  elementos). Si  $q = p^r$  denotamos por  $\mathbb{F}_q$  el cuerpo finito de  $q$  elementos.

**Ejemplo 2.** Construir un cuerpo con 4 elementos.

Consideremos  $\mathbb{F}_2[x]$  y el ideal  $\langle x^2 + x + 1 \rangle$ , claramente  $\langle x^2 + x + 1 \rangle$  es irreducible en  $\mathbb{F}_2[x]$ , además  $\mathbb{F}_2[x]$  es D.I.P., no existe ningún ideal  $I$ ,  $\langle x^2 + x + 1 \rangle \subset I \subset \mathbb{F}_2[x]$ , por tanto  $\langle x^2 + x + 1 \rangle$  es maximal, de aquí, se tiene que  $\mathbb{F}_2[x]/\langle x^2 + x + 1 \rangle$  es cuerpo. Hacemos  $\alpha = x + (x^2 + x + 1)$ . Los elementos de  $\mathbb{F}_2[x]/\langle x^2 + x + 1 \rangle$  son:

$$\begin{aligned}\alpha &= x + (x^2 + x + 1) \\ \alpha + 1 &= (x + 1) + (x^2 + x + 1) \\ 1 &= 1 + (x^2 + x + 1) \\ 0 &= 0 + (x^2 + x + 1)\end{aligned}$$

$\text{card}(\mathbb{F}_2[x]/\langle x^2 + x + 1 \rangle) = 4$  de la observación anterior.

**Algoritmo 1** (Algoritmo de Euclides). Se define el **algoritmo de Euclides** para calcular **máximo**

*común divisor*, de forma que sean  $a, b \in \mathbb{N}$ , con  $a \geq b$  hacemos:

$$\begin{aligned} a &= c_1 b + r_1 \\ b &= c_2 r_1 + r_2 \\ r_1 &= c_3 r_2 + r_3 \\ r_2 &= c_4 r_3 + r_4 \\ &\dots \\ r_{s-2} &= c_s + r_{s-1} + r_s \\ r_{s-1} &= c_{s+1} + r_s \end{aligned}$$

Podemos obtener  $r_s = \text{mcd}(a, b)$  del modo siguiente:

$$\begin{aligned} r_1 &= a - c_1 b \\ b &= c_2(a - c_1 b) + r_2 \\ r_2 &= -c_2 a + (1 + c_1)b \end{aligned}$$

De este modo, podemos obtener  $r_s$  como continuación de  $a, b$ , de modo:  
 $\text{mcd}(a, b) = r_s = \lambda a + \mu b$

**Observación 1.**  $\lambda, \mu$  se obtienen de modo efectivo a partir de las divisiones anteriores.

**Ejemplo 3.** Calcular  $\text{mcd}(139, 20)$ .

Y además,  $\lambda, \mu \in \mathbb{Z}$  tales que  $\lambda 139 + \mu 20 = \text{mcd}(139, 20)$ .

Aplicando el algoritmo de Euclides:

$$\begin{aligned} 139 &= 6 \cdot 20 + 19 \\ 20 &= 1 \cdot 19 + 1 \\ 19 &= 139 - 6 \cdot 20 \\ 1 &= 20 - 1 \cdot 19 = 20 - 1(139 - 6 \cdot 20) = 7 \cdot 20 - 1 \cdot 139 \end{aligned}$$

**Observación 2.** El inverso de 20 en  $\mathbb{F}_{139}$ :

$$7 \cdot 20 - 1 \cdot 139 = 1.$$

Si reducimos mód 139:

$$7 \cdot 20 = 1 \quad \text{mód } 139$$

El inverso de 20 en  $\mathbb{F}_{139}$  es 7

## Capítulo 2

# Códigos autocorrectores

### 2.1. Parámetros de un código

**Definición 3** (Alfabeto finito). Decimos que  $\mathcal{A}$  es un **alfabeto finito** si  $\mathcal{A}$  es conjunto finito de  $q$  símbolos.

**Definición 4** (Código). Decimos que  $\mathcal{C}$  es un **código** si  $\mathcal{C} \subset \mathcal{A}^n$ ,  $\mathcal{C} \neq \emptyset$

**Definición 5** (Palabra). Decimos que  $x$  es una **palabra** si  $x \in \mathcal{A}$ .

**Definición 6** (Palabra-código). Decimos que  $c$  es una **palabra-código** si  $c \in \mathcal{C}$ .

**Ejemplo 4.**  $\mathcal{A} = \mathbb{F}_2$ . Tomamos  $\mathcal{C} = \mathcal{A}$ , aquí hay más probabilidad de error.  $\mathcal{C}_2 = \{(000), (111)\} \subset \mathbb{F}_2^3$  código.

**Definición 7** (Parámetros de un código). Los **parámetros de un código** son:

- **Longitud:**  $n$ .
- **Razón de información:**  $\frac{\log_q \#\mathcal{C}}{n}$
- **Distancia de Hamming:** Sean  $x = (x_1, x_2, \dots, x_n) \in \mathcal{A}^n$ ,  $y = (y_1, y_2, \dots, y_n) \in \mathcal{A}^n$  se define  $d(x, y) = \#\{i : x_i \neq y_i\}$ . Define una distancia métrica.
- **Distancia mínima:**  $d(\mathcal{C}) = \min\{d(c, c') : c, c' \in \mathcal{C}, c' \neq c\}$

**Observación 3.** Sea  $\mathcal{C}$  un código con distancia mínima  $2d(\mathcal{C}) + 1$ . Además:

$$d(x, c_1) \leq d(\mathcal{C})$$

$$d(x, c_2) \leq d(\mathcal{C})$$

Ahora, por la desigualdad triangular

$$d(c_1, c_2) \leq d(c_1, x) + d(c_2, x) \leq d(\mathcal{C}) + d(\mathcal{C})$$

Por tanto, el código se decodificará con mayor facilidad.

**Observación 4.** Si  $d(\mathcal{C}) = 2d(\mathcal{C}) + 1$ , dos bolas cualesquiera  $B(c_1, d(\mathcal{C}))$ ,  $B(c_2, d(\mathcal{C}))$ ,  $c_1 \neq c_2 \in \mathcal{C}$  son disjuntas.

**Definición 8** (Código perfecto). *Un **código perfecto** es un código con  $d(\mathcal{C}) = 2d(\mathcal{C}) + 1$  y tal que  $\{B(c, d(\mathcal{C})) : c \in \mathcal{C}\} = \mathcal{A}^n$*

El **Ejemplo 4** corresponde a un código perfecto.

**Notación 1.** *Escribimos  $[n, M, d]$  – código si queremos representar un código de longitud  $n$ ,  $M$  elementos y diistancia mínima  $d$ .*

**Ejemplo 5.** *Sea  $\mathcal{C}_1 = \{0, 1\}$  y  $\mathcal{C}_2 = \{(000), (111)\} \subset \mathbb{F}_2^3$ .*

*Sea  $0 < p < \frac{1}{2}$  la probabilidad de que se produzca un error al trasmitir un bit.*

*Como  $p < \frac{1}{2}$  proponemos un algoritmo de decodificación, por **máximo verosimilitud**, lo más probable es que la palabra enviada del código sea la palabra de las del código que está más cercana a la palabra recibida.*

*Para  $\mathcal{C}_1$ , la probabilidad de decodificar correctamente la palabra recibida es  $1 - p$ .*

*Para  $\mathcal{C}_2$ : (000) la decodificamos correctamente si al enviarla nos llega (000), la probabilidad de que sea correcta sería  $(1 - p)^3$ , (100), la probabilidad de que sea correcta sería  $p(1 - p)^2$ , (010), la probabilidad de que sea correcta sería  $p(1 - p)^2$  y (001) la probabilidad de que sea correcta sería  $p(1 - p)^2$ .*

*Por tanto, la probabilidad sería  $(1 - p)^3 + 3p(1 - p)^2$ .*

*Si tomamos  $p = 0,1$ , tendríamos para  $\mathcal{C}_1 : 0,9$  y para  $\mathcal{C}_2 : 0,972$ .*

**Ejemplo 6.** *Sea  $\mathcal{C} = \{(000), (111)\} \subset \mathbb{F}_2^3$ .*

Recibida	Decodificar
(000)	(100)
(100)	(100)
(110)	(100)
(101)	(100)
(010)	(100)
(001)	(100)
(011)	(111)
(111)	(111)

*De los que decodificamos como (100) tenemos una probabilidad de acierto de  $(1 - p)^3 + 3p(1 - p)^2 + 2p^2(1 - p)$ , en el otro caso tenemos  $(1 - p)^3 + p(1 - p)^2$ .*

## 2.2. Códigos lineales

**Definición 9** (Código lineal). *Un **código lineal**  $\mathcal{C} \subset \mathbb{F}_q^n$  donde  $\mathcal{C}$  es un subespacio vectorial de  $\mathbb{F}_q^n$*

Gracias a que es una estructura lineal, el código podemos simplificarlo teniendo en cuenta la base del subespacio vectorial.

**Definición 10** (Peso de  $x$ ). *Sea  $x = (x_1, x_2, \dots, x_n) \in \mathbb{F}_q^n$ . El **peso** de  $x$  es  $w(x) = \#\{i : x_i \neq 0\}$*

**Notación 2.**  $\mathcal{C} \subset \mathbb{F}_q^n$ . Se define  $W(\mathcal{C}) = \min\{w(x) : x \in \mathcal{C} / \{0\}\}$



**Proposición 1.** Si  $\mathcal{C}$  es lineal, entonces  $d(\mathcal{C}) = W(\mathcal{C})$

*Demostración.* Sean  $c_1 \neq c_2 \in \mathcal{C}$  que cumplen que  $d(c_1, c_2) = d(\mathcal{C})$

Dado que  $d(\mathcal{C}) = d(c_1, c_2) = w(c_1 - c_2)$ , pero dado que  $c_1 - c_2 \neq 0$ , tenemos que  $w(c_1 - c_2) \geq W(\mathcal{C})$ , por tanto,  $d(\mathcal{C}) \geq W(\mathcal{C})$ .

Sea  $c \in \mathcal{C} - \{0\}$ , tal que cumple que  $w(c) = W(\mathcal{C})$ .

$W(\mathcal{C}) = w(c) = w(c - 0) = d(c, 0) \geq d(\mathcal{C})$  □

**Notación 3.** Si  $\mathcal{C} \subset \mathbb{F}_q^n$  lineal,  $\dim_{\mathbb{F}_1} \mathcal{C} = k$  decimos que  $\mathcal{C}$  es un  $[n, k]$  - código.

**Definición 11** (Matriz generadora). **Matriz generadora** de un  $[n, k]$  - código es una matriz  $G$ , con dimensiones  $k \times n$  cuyas filas forman una base de  $\mathcal{C}$

**Ejemplo 7.** Si  $\mathcal{C} = \{(000), (111)\}$ , la matriz generadora es  $(111)$

**Definición 12** (Matriz de control de paridad). **Matriz de control de paridad** de un  $[n, k]$  - código lineal es una matriz  $H$  de orden  $(n - k) \times n$  tal que:

$$\mathcal{C} = \{x \in \mathbb{F}_q^n : xH^t = 0\}$$

**Observación 5.** Si  $\mathcal{G} = (I_k | P)$  es generadora, con  $P$  con  $n - k$  columnas, entonces  $H = (-P^t | I_{n-k})$  es de control de paridad.

**Ejemplo 8.** En  $\mathbb{F}_3^5$  consideramos el código con matriz generadora:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Tratemos de dar una base de la matriz:

$$\begin{aligned} \mathcal{C} &= \{\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 : \lambda_i \in \mathbb{F}_3\} \\ &= \{(\lambda_1, \lambda_2, \lambda_3, \lambda_1 + \lambda_2, \lambda_1 + \lambda_3) : \lambda_i \in \mathbb{F}_3\} = \\ &= \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}_3^5 : (-x_3 - x_2 + x_4 = 0, x_4 = x_1 + x_2, x_5 = x_1 + x_3)\} = \end{aligned}$$

La matriz de las ecuaciones anterior es:

$$\begin{pmatrix} -1 & -1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = (0, 0)$$

La transpuesta de la matriz anterior es la matriz de control de paridad de  $\mathcal{C}$

**Algoritmo 2** (Algoritmo de decodificación usando el síndrome). Supongamos que  $\mathcal{C} \subset \mathbb{F}_q^n$  es un código lineal, sea  $x \in \mathbb{F}_q^n$ , podemos decodificar  $x$  usando el código  $\mathcal{C}$  comparando  $x$  con cada una de las palabras del código.

Consideramos:

$$x - \mathcal{C} = \{x - c : c \in \mathcal{C}\} = x + \mathcal{C}$$

$x + \mathcal{C}$  es el conjunto de elementos que están en la misma clase  $\mathbb{F}_q^n / \mathcal{C}$ .

En el conjunto  $x - \mathcal{C}$  consideramos un elemento  $e$  que tenga peso mínimo.

Si  $e \in \mathcal{C}$  será de la forma  $e = x - c$ . Entonces  $c \in \mathcal{C}$  está a la menor distancia de la palabra  $x$ .

Podemos decodificar la palabra  $x$  por la palabra  $c$ , que es  $c = x - e$ .

A la palabra  $e$  se le llama **vector error** (asociado a  $x$ ).

**Definición 13** (Síndrome de  $x$ ). Sea  $H$  la matriz de control de paridad del código  $\mathcal{C}$ , consideramos  $x \in \mathbb{F}_q^n$ , se define **el síndrome** de  $x$  como  $xH \in \mathbb{F}_q^{n-k}$ .

**Observación 6.** Dos palabras  $x, x' \in \mathbb{F}_q^n$  tienen el mismo síndrome si y sólo si  $x + \mathcal{C} = x' + \mathcal{C}$ . En efecto, si  $x + \mathcal{C} = x' + \mathcal{C} \iff xH^t - x'H^t = 0 \iff xH^t = x'H^t$

**Algoritmo 3** (Algoritmo de decodificación usando síndrome). A continuación se muestra el **algoritmo de decodificación usando síndrome**

1. Para cada clase  $x + \mathcal{C}$  de  $\mathbb{F}_q^n/\mathcal{C}$  consideramos un  $e \in x + \mathcal{C}$  de peso mínimo.
2. Para cada clase  $x + \mathcal{C}$  calculamos el síndrome  $xH^t$ . Asociamos a ese síndrome el vector error  $e$ .
3. Cuando recibimos una palabra  $x$ , calculamos su síndrome, que tiene asociado un vector error  $e$ , y decodificamos  $x$  por  $c = x - e$ .

**Ejemplo 9.** En  $\mathbb{F}_2^5$  consideramos el código  $\mathcal{C}$  con matriz generadora:

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Para calcular  $\mathbb{F}_2^5/\mathcal{C}$ , ampliamos la matriz  $G$  de modo que sea una base de  $\mathbb{F}_2^5$ , de este modo la base de  $\mathbb{F}_2^5/\mathcal{C}$  sería lo que se 'amplia'.

Es fácil ver que entonces que  $\{(00010) + \mathcal{C}, (00001) + \mathcal{C}\}$  es una base de  $\mathbb{F}_2^5/\mathcal{C}$ .

De esta definición, las distintas clases de  $\mathbb{F}_2^5/\mathcal{C}$  son:

1.  $(00000) + \mathcal{C}$
2.  $(00010) + \mathcal{C}$
3.  $(00001) + \mathcal{C}$
4.  $(00011) + \mathcal{C}$

En el caso 1, dado que el  $(00000)$  está en el código, es fácil ver que el vector de error es  $e = (00000)$ .

Por otro lado, al sumar a 2 ella misma, tenemos  $(00000)$ , por tanto,  $e = (00010)$ .

Para 3 es análogo,  $e = (00001)$ .

Finalmente, para el último tenemos  $e = (10000)$ .

Calculamos ahora los síndromes para cada clase; empezamos calculando una matriz de control de paridad  $H$  del código  $\mathcal{C}$ , transponiendo  $G$ , tenemos:

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Ahora, para cada clase calculamos  $xH^t$ .

1.  $(0, 0) \longleftrightarrow e = (00000)$
2.  $(1, 0) \longleftrightarrow e = (00010)$
3.  $(0, 1) \longleftrightarrow e = (00001)$

$$4. (1, 1) \longleftrightarrow e = (10000)$$

*Se puede ver que las últimas componentes de los vectores  $x$  coinciden con el valor del síndrome, esto no es una casualidad.*

Ahora tomamos cualquier palabra  $x = (01101)$  calculamos su síndrome:

$$xH^t = (1, 0) \longleftrightarrow e = (00010)$$

Entonces,

$$c = x - e = (01101) - (00010) = (01111)$$

### 2.2.1. Códigos de Hamming

Supongamos que estamos  $\mathbb{F}_q^r$ , y queremos un conjunto maximal de vectores de  $\mathbb{F}_q^r$ , con la condición de que cualesquiera dos vectores del conjunto sean linealmente independientes. En  $\mathbb{F}_q^r - \{0\}$ , consideramos la relación de equivalencia:

$$(b_1, \dots, b_r) \equiv (a_1, \dots, a_r) \iff (b_1, \dots, b_r) = \lambda(a_1, \dots, a_r), \lambda \in \mathbb{F}_q$$

Cada clase de equivalencia tiene  $q - 1$  elemento  $= (\#(\mathbb{F}_q - \{0\}))$ .

Ahora,  $\# \{\mathbb{F}_q - \{0\}\} = q^r - 1$ .

Luego

$$\# \mathbb{F}_q^r - \{0\} / \sim = \frac{q^r - 1}{q - 1}$$

## 2.3. Algunos códigos buenos

## 2.4. Códigos cíclicos



## Capítulo 3

# Criptografía

3.1. Criptosistemas simétricos

3.2. Criptosistemas de clave pública



# Autoría

Estas notas se han realizado en base a los apuntes de clase del profesor **Bartolomé López Jiménez** de la Universidad de Cádiz.