Seminario del Profesorado en Matemática

Introducción a la Teoría de Códigos

por Emmanuel Mauricio López

Universidad Nacional de Salta
Facultad de Ciencias Exactas
Departamento de Matemática

Director: Lic. Gonzalo Maximiliano López

Resumen

En los últimos 70 años matemáticos e ingenieros pusieron sus esfuerzos para garantizar la transmisión de información de manera confiable y eficiente. Estos esfuerzos dieron forma a la Teoría de Códigos, que fue más allá del problema de la transmisión de información, teniendo aplicaciones en, por ejemplo, el almacenamiento de datos. Los códigos de bloques son una solución al problema de la transmisión de información debido a que mediante ellos se puede detectar y corregir errores ocurridos durante la transmisión.
Una clase especial de códigos de bloque son los códigos lineales y los códigos cíclicos y son éstos los que exponemos en este trabajo, detallamos su codificación, decodificación y su capacidad de detectar y corregir errores.
Encontrar códigos capaces de codificar la mayor cantidad de mensajes posibles y que además sean buenos códigos, en cuanto a su capacidad de detectar y corregir errores, es el principal problema de la Teoría de Códigos. Aunque determinar esa cantidad máxima es un problema abierto, es posible acotarla. En este trabajo mostramos de forma breve cotas superiores e inferiores y dirigimos nuestra atención a una, la cota de Hamming, y con ella definimos a los códigos perfectos los cuáles ejemplificamos.

Agradecimientos

Quiero expresar mis agradecimientos a Gonzalo, mi director, por su disposición, dedicación y la paciencia en la elaboración de este trabajo. Trabajar junto a él ha sido una experiencia gratificante, sus conocimientos y experiencias han contribuido significativamente en mi formación académica.

A mi familia por el apoyo y aliento constante a que cumpla mis metas. En especial a mis padres por el amor y el cariño recibido a diario y por la paciencia durante mis años de estudio, por compartir la alegría en mis aciertos y brindarme su consuelo en las caídas. Papá, Mamá, les dedico este trabajo que es también fruto de sus esfuerzos.

Agradezco a mis amigos por su permanente compañia. En mi primer año en la universidad conocí a Pamela y a Antonio, y desde ese momento hemos sido grandes amigos. Les agradezco su amistad y el haber motivado mi interés a la matemática de una manera profunda como así también el constante aliento a dar siempre lo mejor de mi en cada uno de mis proyectos. Recuerdo con mucha alegría los momentos compartidos.

Y a la Universidad Nacional de Salta junto a todos los docentes de quienes he recibido formación y que buscan siempre la excelencia en lo académico y lo humano.

Índice general

Introducción

La teoría de códigos, junto con la teoría de la información nacen de una publicación realizada por Claude Shannon (ingeniero electrónico y matemático estadounidense) en 1948 titulada *Una teoría matemática de la comunicación* [1]. En esas décadas, Shannon trabajaba en los Laboratorios Bell en Estados Unidos, y centró toda su atención en los procesos de transmisión de información a través de canales de comunicación ruidosos. El aporte revolucionario que hizo fue cuantificar la información transmitida y demostrar la existencia de códigos los suficientemente buenos para proteger el mensaje del ruido o interferencia presente en los canales de comunicación.
Shannon no deja un método o una receta para encontrar estos códigos, sólo demuestra su existencia [2], sin embargo provocó la busqueda, por parte de muchos matemáticos e ingenieros, de tales códigos, dándole forma a la teoría de códigos.
Almacenamiento de datos, comunicación satelital, grabación de discos compactos, son algunas de las aplicaciones de la teoría de códigos. Son muchas las áreas de la matemática que son necesarias para el estudio de esta teoría, la principal es el álgebra, pero también involucra conceptos de probabilidad, de combinatoria y teoría de números.
Una de las primeras palabras en la publicación de Shannon en 1948 fueron *El problema fundamental de la comunicación es el de reproducir en un punto exactamente o aproximadamente un mensaje seleccionado en otro punto*, y propone un modelo de comunicación (Fig. (0.1)).

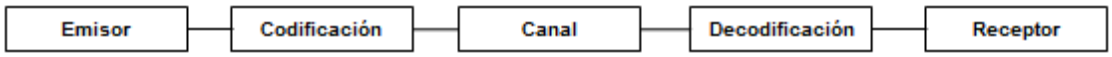


Figura 0.1.: Modelo matemático o telegráfico de la comunicación

Este trabajo se exponen los conceptos básicos de la Teoría de códigos, como lo hace la literatura clásica, y va dirigido al lector que tiene conocimientos en Estructuras Algebraicas. El lector descubrirá en los capítulos que siguen que lo expuesto es una aplicación directa a problemas prácticos de conceptos pertenecientes a Estructuras Algebraicas y al Algebra Lineal principalmente.
En la mayoría del trabajo estudiamos a los códigos lineales, su codificación y decodificación en general, pero también discutimos sobre el principal problema de esta Teoría que consiste en la busqueda de códigos óptimos. Presentamos a la familia de

códigos Hamming y Golay, códigos lineales que nos ayudan a ejemplificar la capacidad que tiene un código de corregir y detectar errores, la existencia de técnicas de decodificación eficientes para códigos lineales, y para ejemplificar un tipo de código lineal llamado código cíclico, cuyo estudio lo hacemos utilizando principalmente polinomios sobre un cuerpo finito.

1. Códigos y métodos de decodificación

Como dijimos con anterioridad, estamos interesados en códigos detectores y correctores de errores. En este primer capítulo introduciremos el concepto de código, enunciaremos de forma general dos métodos de decodificación de los cuáles se forman algoritmos de decodificación para algunos códigos especiales que presentaremos en los siguientes capítulos. Introduciremos una métrica llamada de Hamming, y la usaremos para describir a los códigos en general y establecer condiciones para identificar a aquellos que detectan y corrigen errores.

1.1. Primeros conceptos

Una primer idea sobre la tarea de cómo transmitir un mensaje es presentada en el trabajo de Shannon (ver Figura (0.1)). Pero ¿qué tipo de mensaje se transmiten? Transmitimos cadenas finitas de elementos que pertenecen a un conjunto finito al que llamaremos alfabeto. Así, por ejemplo, cada palabra que estamos escribiendo son cadenas finitas escritas con las 26 letras del alfabeto castellano. Formalicemos algunas conceptos.

Definición 1.1.1. Sea $A = \{a_1, a_2, ..., a_q\}$ un conjunto con q elementos y A^n el conjunto de las n-uplas sobre A, con n un entero positivo. Llamaremos a A alfabeto y a sus elementos símbolos.

(a) Sea $x \in A^n$. Diremos que x es una *palabra* de longitud n sobre A o escrita con el alfabeto A.

(b) Un *código de bloques* de longitud n sobre A es un conjunto C no vacío de palabras, de la misma longitud n escritos con el alfabeto A.

(c) Un elemento de C se denomina *palabra código* de C.

(d) El número de palabras códigos de C, denotada por $|C|$, es llamado *tamaño* de C.

(e) El *radio de información* del código C de longitud n está definido por $(log_q |C|)/n$.

(f) Un código de longitud n y tamaño M es llamado un *código (n,M).*

Observación 1.1.2.

1. En este trabajo, y como es usual en la práctica, los alfabetos que usaremos serán cuerpos finitos de orden q denotados por F_q, donde q es la potencia de un número primo. El conjunto de palabras de longitud n escritas sobre F_q se denota F_q^n.
2. $x \in A^n$ en textos de álgebra suele estar escrito como $x = (x_1, x_2, ..., x_n)$ con $x_1, x_2, ..., x_n \in A$. Preferimos aquí escribir x como $x = x_1x_2...x_n$.
3. Un código de bloque de longitud n también se dice que es un código de longitud n.
4. La cantidad de palabras de una longitud fija que se pueden formar con símbolos de un alfabeto, es una cantidad finita. Por ejemplo, si A tiene q elementos entonces la cantidad de palabras de longitud n que se pueden formar es q^n.

Ejemplo 1.1.3. El alfabeto $F_2 = \{0, 1\}$ es un cuerpo con las operaciones suma y producto módulo 2. Un código sobre F_2 se llama *código binario*. Algunos ejemplos de ellos son:

(a) $C_1 = \{00, 01, 10, 11\}$ es un código (2,4)

(b) $C_2 = \{0011, 0101, 1010, 1100, 1001, 0110\}$ es un código (4,6)

(c) $C_3 = \{00000, 01101, 10110, 11011\}$ es un código (5,4)

Otros cuerpos que usaremos más frecuentemente son: $F_3 = \{0, 1, 2\}$ (llamado cuerpo ternario), cuyas operaciones son la suma y el producto módulo 3, y el cuerpo $F_4 = \{0, 1, w, \bar{w}\}$ (llamado cuerpo cuaternario), cuyas operaciones están definidas por las siguientes tablas [3]:

$+$	0	1	w	$\bar{w}$
0	0	1	w	$\bar{w}$
1	1	0	$\bar{w}$	w
w	w	$\bar{w}$	0	1
$\bar{w}$	$\bar{w}$	w	1	0

$\cdot$	0	1	w	$\bar{w}$
0	0	0	0	0
1	0	1	w	$\bar{w}$
w	0	w	$\bar{w}$	1
$\bar{w}$	0	$\bar{w}$	1	w

Diremos que una palabra código es enviada o es transmitida haciendo referencia a la etapa en que la palabra código pasa a través del canal de comunicación para su posterior decodificación. También diremos que una palabra código es retransmitida, o se hace una retransmisión, cuando luego de ser transmitida, volvemos a transmitirla.
Las palabras código se transmiten dígito por dígito. Así, para transmitir una palabra de longitud n de un código C, se deben hacer n transmisiones, una por cada dígito.

Los mensajes son en principio cualquier palabra con o sin sentido, de longitud variable que puede no estar escrita con el mismo alfabeto que el código que se usará para codificar. En el Capítulo 2 sin embargo, codificaremos mensajes de una misma

longitud escritos con el mismo alfabeto que el código que se usará para codificar. Este tipo de mensaje se formalizará en el capitulo siguiente.
Entendemos la codificación como una correspondencia uno a uno entre los mensajes y las palabras códigos de un código C. Por ejemplo, sea C_1 el código binario $C_1 = \{00, 01, 10, 11\}$, entonces podemos codificar los mensajes Norte, Sur, Este y Oeste de la siguiente manera

$$\begin{aligned} \text{Norte} &\longrightarrow 00 \\ \text{Sur} &\longrightarrow 01 \\ \text{Este} &\longrightarrow 10 \\ \text{Oeste} &\longrightarrow 11 \end{aligned}$$

Pasemos a definir lo que es el medio por el cual es transmitida la información, el *canal de comunicación* [4].

Definición 1.1.4. Un *canal de comunicación* consiste en un alfabeto $A = \{a_1, a_2, ..., a_q\}$ junto con las probabilidades $P(a_j \text{ recibido} \mid a_i \text{ enviado})$ que satisfacen

$$\sum_{j=1}^{q} P(a_j \text{ recibido} \mid a_i \text{ enviado}) = 1 \text{ para todo } i = 1, ..., q$$

donde $P(a_j \text{ recibido} \mid a_i \text{ enviado})$ denota la probabilidad condicional de que a_j es recibido, dado que a_i es enviado. En un canal de comunicación sólo se envían palabras código.

Definición 1.1.5. Un canal de comunicación es llamado *sin memoria* si el resultado de una transmisión es independiente al resultado de una transmisión previa, en otras palabras, si $y = y_1y_2...y_n$ y $x = x_1x_2...x_n$ son palabras de longitud n entonces

$$P(y \text{ recibido} \mid x \text{ enviado}) = \prod_{i=1}^{n} P(y_i \text{ recibido} \mid x_i \text{ enviado}) \tag{1.1.1}$$

Definición 1.1.6. Un *canal simétrico q-ario* es un canal de comunicación sin memoria que tiene un alfabeto de tamaño q tal que

(a) cada símbolo transmitido tienen la misma probabilidad p (con $p < 1/2$) de ser recibido con error.

(b) si un símbolo es recibido con error, entonces cada uno de los $q - 1$ posibles errores son igualmente probables.

En particular el *canal binario simétrico* es un canal sin memoria sobre el alfabeto F_2 y tiene la siguiente probabilidad

$$\begin{aligned} P(1 \text{ recibido} \mid 0 \text{ enviado}) &= P(0 \text{ recibido} \mid 1 \text{ enviado}) = p \\ P(1 \text{ recibido} \mid 1 \text{ enviado}) &= P(0 \text{ recibido} \mid 0 \text{ enviado}) = 1 - p \end{aligned}$$

donde p es la probabilidad de error en la transmisión del símbolo.

Ejemplo 1.1.7. Sea $C = \{000, 111\}$ un código. Supongamos que enviamos una palabra código sobre el canal binario simétrico con probabilidad de error $p = 0{,}04$ y la palabra recibida es 110. Por (1.1.1) se tiene que:

$$\begin{aligned} P(\text{110 recibido} \mid \text{000 enviado}) &= P(\text{1 recibido} \mid \text{0 enviado})^2 \times P(\text{0 recibido} \mid \text{0 enviado}) \\ &= (0{,}04)^2(0{,}96) = 0{,}001536 \end{aligned}$$

$$\begin{aligned} P(\text{110 recibido} \mid \text{111 enviado}) &= P(\text{1 recibido} \mid \text{1 enviado})^2 \times P(\text{0 recibido} \mid \text{1 enviado}) \\ &= (0{,}96)^2(0{,}04) = 0{,}036864 \end{aligned}$$

y asi es más probable que la palabra código 111 haya sido enviada.

De forma general, supongamos que se envía una palabra código de C, un código de longitud n, y que han ocurrido r errores, con $0 \leq r \leq n$. Supongamos que la probabilidad de error es p, con $p < 1/2$. Si se recibe la palabra y, entonces se tiene que

$$P(y|c) = p^r(1-p)^{n-r} \tag{1.1.2}$$

Sabemos que por un canal de comunicación solo se transmiten las palabras códigos. Supongamos que la palabra x es recibida. Si x es una palabra código, afirmamos que no hubo error en la transmisión. Si no lo es, han ocurrido errores en el envío, i.e. hay coordenadas de la palabra código enviada que han cambiado de valor. Así, necesitamos métodos de decodificación. En este trabajo enunciaremos dos métodos generales de decodificación: la decodificación máxima más probable (maximun likelihood decoding) y la decodificación de la distancia mínima (nearest neighbour decoding).

1.2. Métodos de decodificación

Previo a describir los métodos de decodificación, intoduciremos lo que es uno de los conceptos más importantes en la teoría de códigos, la *distancia Hamming*, una contribución del matemático estadounidense Richard Hamming (1915-1998).

Definición 1.2.1. Sean $x = x_1x_2...x_n$, $y = y_1y_2...y_n \in F_q^n$. La distancia (Hamming) $d(x, y)$ entre x e y es número de coordenadas en las cuales x e y difieren. Es decir:

$$d(x, y) = |\{i|1 \leq i \leq n,\, x_i \neq y_i\}|$$

Ejemplo 1.2.2. (a) Sean F_2^6 y $x = 010111$, $y = 111010$, $z = 010101$ entonces:

$$d(x,y) = 4$$
$$d(y,z) = 5$$
$$d(x,z) = 1$$

(b) Sean F_4^3 y $x = 01w$, $y = w1\bar{w}$, $z = \bar{w}\bar{w}\bar{w}$ entonces:

$$d(x,y) = 2$$
$$d(y,z) = 2$$
$$d(x,z) = 3$$

Teorema 1.2.3. *La distancia Hamming es una función distancia*

Demostración. Debemos probar que la distancia Hamming satisface las siguientes condiciones. Sean $x, y, z \in F_q^n$

1. $d(x,y) \geq 0$
2. $d(x,y) = 0$ si y solo si $x = y$
3. $d(x,y) = d(y,x)$
4. $d(x,y) \leq d(x,z) + d(z,y)$

Si $x = x_1x_2...x_n$, $y = y_1y_2...y_n$ y $z = z_1z_2...z_n$

1. Se tiene que $d(x,y)$ es no negativo por definición de cardinal de un conjunto, así la distancia Hamming cumple la primera condición.
2. $\Rightarrow$) Supongamos que $x \neq y$. Luego existe un número natural i con $0 < i \leq n$ tal que $x_i \neq y_i$. Así $d(x,y) > 0$. Absurdo, pues $d(x,y) = 0$. Finalmente se tiene que $x = y$.
   $\Leftarrow$) $x = y$ entonces $x_i = y_i$ para cualquier i.
   Así $d(x,y) = |\{i|\ 1 \leq i \leq n,\ x_i \neq y_i\}| = 0$.
3. $$\begin{aligned} d(x,y) &= |\{i|1 \leq i \leq n \quad x_i \neq y_i\}| \\ &= |\{i|i \leq 1 \leq n \quad y_i \neq x_i\}| \\ &= d(y,x). \end{aligned}$$
4. Sea $U = \{i|x_i \neq y_i\}$. Así $d(x,y) = |U|$.
   Sea $S = \{i|x_i \neq y_i \wedge x_i = z_i\}$ y $T = \{i|x_i \neq y_i \wedge x_i \neq z_i\}$
   Luego U es la unión disjunta de S y T por lo que
   $$|U| = |S| + |T| \tag{1.2.1}$$
   Como $d(x,z) = |\ \{i|x_i \neq z_i\}\ |$ se tiene que $|T| \leq d(x,z)$.
   Por otro lado si $i \in S, x_i \neq y_i \wedge x_i = z_i$ por lo que $y_i \neq z_i$ y $|S| \leq d(z,y)$
   Luego por (1.2.1) se tiene que $d(x,y) \leq d(x,z) + d(z,y)$.

□

A continuación presentamos los dos métodos de decodificación mecionados arriba que usaremos para determinar qué palabra código c fue enviada a partir de una palabra y recibida.
Supongamos que se han enviado palabras códigos de un código C sobre un canal de comunicación. Si recibimos una palabra y, calculamos para cada $c \in C$ la probabilidad

$$P(y \text{ recibido} \mid c \text{ enviado}) \tag{1.2.2}$$

La *decodificación máxima más probable* consiste en maximizar (1.2.2) i.e encontrar c_y tal que

$$P(y \text{ recibido} \mid c_y \text{ enviado}) = \max_{c \in C} P(y \text{ recibido} \mid c \text{ enviado})$$

Hay dos clases de decodificación máxima más probable:

(a) *Decodificación máxima más probable completa.* Supongamos que la palabra y es recibida, luego encontramos $c_y \in C$ que maximiza (1.2.2). Si hay más de una palabra código que maximiza (1.2.2), elegimos una arbitrariamente.

(b) *Decodificación máxima más probable incompleta.* Supongamos que la palabra y es recibida, luego encontramos $c_y \in C$ que maximiza (1.2.2). Si hay más de una palabra código que maximiza (1.2.2), hacemos una retransmisión.

Volvamos al supuesto que se han enviado palabras códigos de un código C sobre un canal de comunicación y que hemos recibido la palabra y.
El segundo método de decodificación, *la decodificación de la distancia mínima o del objeto más cercano* consiste en encontrar la palabra código $c_y \in C$ tal que

$$d(y, c_y) = \min_{c \in C} d(y, c)$$

Al igual que la decodificación máxima más probable, hay dos tipos de la decodificación de la distancia mínima:

(a) *Decodificación de la distancia mínima completa.* Supongamos que la palabra y es recibida, luego encontramos $c_y \in C$ tal que $d(y, c_y) = \min_{c \in C} d(y, c)$. Si hay más de una palabra código que cumple con la igualdad anterior, elegimos una arbitrariamente.

(b) *Decodificación de la distancia mínima incompleta.* Supongamos que la palabra y es recibida, luego encontramos $c_y \in C$ tal que $d(y, c_y) = \min_{c \in C} d(y, c)$. Si hay más de una palabra código que cumple con la igualdad anterior, hacemos una retransmisión.

Supongamos que la probabilidad de error en la transmisión de un símbolo del alfabeto F_q es $p < 1/2$, esto nos dice que

$$P(y_i|c_i) = p \qquad si\ y_i \neq c_i$$

$$P(y_i|c_i) = 1 - p \quad si\ y_i = c_i$$

Suponiendo que estamos trabajando con palabras en F_q^n entonces, por (1.1.2), tenemos que

$$P(y|c) = p^{d(y,c)}(1-p)^{n-d(y,c)} = (1-p)^n \left(\frac{p}{1-p}\right)^{d(y,c)}$$

y al ser $p < 1 - p$, $P(y|c)$ será máxima cuando $d(y, c)$ sea mínima [3]. Esto es la prueba del siguiente teorema.

Teorema 1.2.4. *La decodificación máxima más probable y la decodificación de la distancia mínima son equivalentes.*

Nota 1.2.5. La equivalencia de la que hablamos en el teorema anterior se refiere a que realizar la decodificación máxima más probable implica realizar la decodificación de la distancia mínima y recíprocamente, realizar la decodificación de la distancia mínima implica realizar la decodificación máxima más probable.

El siguiente es un ejemplo donde se aplica la decodificación de la distancia mínima.

Ejemplo 1.2.6. Sea $C = \{00000, 01101, 10110, 11011\}$ un código sobre F_2. Ha sido enviada una palabra código y se recibió la palabra $y = 01001$, entonces:

$$\begin{aligned} d(01001, 00000) &= 2 \\ d(01001, 01101) &= 1 \\ d(01001, 10110) &= 5 \\ d(01001, 11011) &= 2 \end{aligned}$$

Entonces concluimos que la palabra enviada fue 01101.

Definición 1.2.7. Para un código C que contiene al menos dos palabras código, la *distancia mínima* de C, denotada por d o $d(C)$ es

$$d = d(C) = min\{d(x, y)|x, y \in C, x \neq y\}$$

Definición 1.2.8. A un código de longitud n, tamaño M y distancia d se lo escribe como un código (n, M, d). Los números n, M, y d se llaman parámetros del código.

Ejemplo 1.2.9. Sea $C = \{00000, 02101, 12211, 21110\}$ un código sobre F_3. Vemos que:

$$\begin{aligned} d(00000, 02101) &= 3 \quad & d(02101, 12211) &= 3 \\ d(00000, 12211) &= 5 \quad & d(02101, 21110) &= 4 \\ d(00000, 21110) &= 4 \quad & d(12211, 21110) &= 4 \end{aligned}$$

por lo que concluimos que $d(C) = 3$ y así C es un $(5, 4, 3)$ código.

1.3. Códigos detectores y correctores de errores

Veamos cómo la distancia mínima de un código también es útil para caracterizar al código en cuanto a su capacidad de corregir y/o detectar errores.

Definición 1.3.1. Sea u un entero positivo. Un código C es *u-error-detector* si a cualquier palabra código que sufra al menos uno pero no más de u errores, la palabra que resulta no es una palabra código. Un código C es *exactamente u-error-detector* si es u-error-detector pero no $(u+1)$-error-detector.

Ejemplo 1.3.2. El código binario $C = \{00000, 00111, 11111\}$ es 1-error-detector ya que cambiando el valor de una coordenada cualquiera en cualquier palabra código, lo que resulta no es una palabra código. En otras palabras

00000 necesita cambiar tres coordenadas para obtener 00111

00000 necesita cambiar cinco coordenadas para obtener 11111

00111 necesita cambiar dos coordenadas para obtener 11111.

C es además exactamente 1-error-detector, ya que si cambiamos las primeras dos coordenadas de 00111 er resultado es la palabra código 11111 (así C no es 2-error-detector).

El siguiente teorema es muy útil para identificar códigos errores-detectores.

Teorema 1.3.3. *Un código C es u-error-detector si y sólo si $d(C) \geq u+1$. En otras palabras, un código con distancia d es exactamente $(d-1)$-error-detector.*

Demostración. $\Rightarrow$) Supongamos que $d(C) < u+1$ i.e. $d(C) \leq u$. Existen $c_1, c_2 \in C$ tal que $d(c_1, c_2) = d(C) \leq u$. Esto quiere decir que si en c_1 ocurren $d(C) \leq u$ errores y se obtiene c_2, por ser C u-error-detector, $c_2 \notin C$. Esto no puede ocurrir puesto que $c_2 \in C$. Así $d(C) \geq u+1$.

$\Leftarrow$) Tenemos que $d(C) \geq u+1$. Sea $c \in C$. Supongamos que en c ocurren a lo sumo u errores obteniendo la palabra x. Luego $1 \leq d(c, x) \leq u < d(C)$. Así $x \notin C$ por lo que C es u-error-detector. □

Definición 1.3.4. Sea t un entero positivo. Un código C es *t-error-corrector* si la decodificación de la distancia mínima es capaz de corregir t o menos errores, suponiendo la decodificación incompleta. Un código C es *exactamente t-error-corrector* si es t-error-corrector pero no $(t+1)$-error-corrector.

Ejemplo 1.3.5. Consideremos el código binario $C = \{000, 111\}$. Usando la decodificación de la distancia mínima se tiene que

- si 000 es enviado y ocurre un error en la transmisión, entonces la palabra recibida (100, 010 o 001) será decodificada como 000,
- si 111 es enviada y ocurre un error en la transmisión, entonces la palabra recibida (110, 101 o 011) será decodificada como 111.

En resumen, si ocurre un error, éste puede ser corregido. Así, C 1-error-corrector. Si al menos dos errores ocurren, por ejemplo si se envía 000 y 011 es recibida, entonces 011 será decodificada como 111 usando la decodificación de la distancia mínima. Así C es exactamente 1-error-corrector.

Teorema 1.3.6. *Un código C es t-error-corrector si y sólo si $d(C) \geq 2t+1$. En otras palabras, un código con distancia d es un código exactamente $\left\lfloor \frac{1}{2}(d-1) \right\rfloor$-error-corrector.*

Demostración. $\Rightarrow$) Supongamos que $d(C) < 2t+1$. Existen $c, c' \in C$ tal que $d(c, c') = d(C) \leq 2t$.
Observemos que si $d(c, c') < t+1$, c puede transformarse en c' si se producen a lo sumo t errores en c. Por hipótesis esto no puede pasar pues C es t-error-corrector. Luego $d(c, c') \geq t+1$.
Sin pérdida de generalidad, podemos asumir que c y c' difieren en las primeras $d = d(C)$ posiciones, donde $t+1 \leq d \leq 2t$. Si consideramos

$$c = c_1 c_2 ... c_n$$
$$c' = c'_1 c'_2 ... c'_n$$

con $c_i \neq c'_i$ para $i = 1, ..., d$ y $c_i = c'_i$ para $i = d+1, ..., n$, podemos formar una palabra x donde

$$x = c'_1 ... c'_t c_{t+1} ... c_d c_{d+1} ... c_n$$

veamos que si se recibe la palabra x se tiene que

$$d(x, c') = d - t \leq t = d(x, c)$$

Concluimos que, suponiendo se envía c y se recibe x (se produjeron t errores), se puede decodificar erroneamente x como c' o se debe pedir una retrasmisión debido a que $d(x, c') \leq d(x, c)$. Esto es absurdo y provino de suponer que $d(C) < 2t+1$. Así $d(C) \geq 2t+1$.

$\Leftarrow$) Supongamos que $c \in C$ fue enviado y x recibido, ocurriendo t o menos errores i.e. $d(c, x) \leq t$. Debemos probar que para cualquier $c' \in C$ con $c \neq c'$ ocurre que

$d(c,x) < d(c',x)$. Por la desigualdad triangular se tiene que

$$\begin{aligned} d(c,c') &\leq d(c,x) + d(c',x) \\ d(c,c') - d(c,x) &\leq d(c',x) \\ 2t + 1 - t &\leq d(c',x) \\ d(c,x) < t + 1 &\leq d(c',x) \\ d(c,x) &< d(c',x) \end{aligned}$$

□

Ejemplo 1.3.7. El código C del ejemplo (1.2.9) es un código 2-error-detector y 1-error-corrector.

1.4. Ejemplos para ilustrar

Habiendo enunciado los conceptos básicos relacionados con los códigos detectores y correctores de errores, analicemos un ejemplo sencillo para ver realmente por qué codificamos [5].

1.4.1. Movimiento remoto

Supongamos que se desea enviar mensajes para mover de forma remota un vehículo pequeño. El vehículo puede moverse de cuatro maneras distintas: hacia arriba (A), hacia abajo (B), a la derecha (C) y a la izquierda (D). Así tenemos cuatro mensajes posibles para enviar, A, B, C, D. Podríamos codificar estos mensajes en un código binario de la siguiente manera

$$C_1 = \begin{cases} 00 = A \\ 01 = B \\ 10 = C \\ 11 = D \end{cases}$$

Notemos que si en la transmisión ocurre un error, una palabra código se transforma en otra, por ejemplo si se envía 00 y ocurre un error en la segunda posición, se recibe 01. Como 01 es una palabra código, la palabra código más cercana a 01 es ella misma, es decir que, aplicando los métodos de decodificación, no podemos detectar que se ha producido un error y se concluirá erroneamente que 01 fue la palabra enviada. Proponemos el siguiente código.

$$C_2 = \begin{cases} 00000 = A \\ 01101 = B \\ 10110 = C \\ 11011 = D \end{cases}$$

Nuevamente supongamos que se envía 00000 y que ocurre un error en la segunda posición recibiendo la palabra 01000. El lector podrá comprobar que 00000 es la palabra código más cercana a 01000. Así concluímos, correctamente, que la palabra código enviada fue 00000.
Vemos que $d(C_2) = 3$ y así por Teorema (1.3.3) y Teorema (1.3.6), C_2 es 2-error-detector y 1-error-corrector. Diremos que C_2 es mejor que C_1 en cuanto a su capacidad de detectar y corregir errores.

1.4.2. Imágenes satelitales

Es común ver fotografías de Marte, Saturno y otros planetas. Las primeras fotografías de los planetas mencionados fueron tomadas por los satélites Mariner y Voyagers. A fin de transmitir estas fotografías a la Tierra, una vez tomadas, se colocaba un reticulado sobre la imagen y a cada recuadro se le asignaba un número binario del 0 al 63, el cual representa el grado de oscuridad de la imagen en el recuadro. Como el número más extenso dentro de la escala dada contiene seis dígitos, a cada recuadro le corresponde un número de seis dígitos. Si el número binario dentro en el rango de la escala dada tiene menos de 6 dígitos entonces se le anteponen tantos ceros como haga falta para que la cantidad de dígitos sea seis. Si se transmiten únicamente estos seis dígitos, los cuales podemos ver como una 6-upla, es muy probable que, debido a las interferencias producidas por los efectos térmicos del espacio, se reciba un mensaje equivocado. Particularmente el satélite Mariner lo que hacía era codificar cada 6-upla en una 32-upla. De esta forma al enviar las 32-uplas, aunque se produzcan errores en la transmisión, la información que se recibe es mucho más que la que se hubiese recibido si solo se hubieran enviado las 6-uplas y esto facilita a la recuperacion de las 6-uplas originales.

En resumen, estos dos primeros ejemplos muestran que lo que hace la codificación es agregar información al mensaje, lo necesario para que al recibir lo enviado uno pueda recuperar el mensaje o advertir que he han producido errores que pueden cambiar el mensaje original. A continuación presentamos algunos códigos conocidos.

1.4.3. El código EAN

El código EAN (European Article Number) consiste en una serie de números que sirven para identificar productos [6]. Existen dos formatos básicos, uno de 13 dígitos y otro de 8, ambos escrito con el alfabeto $\{0, 1, 2, ..., 9\}$. El que se utiliza generalmente es el de 13 dígitos que, suponiendo que es $x_1x_2...x_{13}$, está formado como sigue:

- los primeros dos o tres dígitos corresponden a la identificación del país. Por ejemplo, a Argentina le corresponde los dígitos 779, a España 84, Alemania utiliza los dígitos desde 400 al 440.

- los cinco dígitos siguientes, o cuatro en caso de haber ocupado tres digitos para identificar el país, corresponden a la identificación de la empresa que produjo el producto.

- los cinco dígitos siguientes se corresponden a la identificación del producto, es un número que la empresa productora le asigna a su producto.

- el dígito restante, x_{13}, es un dígito de control y queda determinado por los demás dígitos de forma tal que se satisfaga

$$\sum_{i \equiv 1 \,(mod\, 2)} x_i + 3 \cdot \sum_{i \equiv 0 \,(mod\, 2)} x_i \equiv 0 \,(mod\, 10)$$

Es claro que la palabra cuyos trece dígitos son ceros pertenece al código EAN y se puede verificar que 9010000000000 también lo hace. Así se tiene que este código tiene distancia mínima 2, por lo tanto, puede detectar 1 error.

Este código es usual verlo en los productos de las góndolas de los supermercado o almacenes, los cuales tienen los 13 dígitos acompañado de otro código llamado EAN código de barra, que facilita la lectura de la información por medio de lectores digitales. Este código codifica los dígitos del código EAN en una secuencia binaria los cuales seran representados luego por medio de barras y espacios. Un 1 corresponderá a una barra, el 0 a un espacio y la aparición consecutiva de 1s y 0s corresponderan respectivamente a barras y espacios más anchos. El EAN código de barras se forma

- con 101 del lado izquierdo, que comienza el código

- la codificación binaria de los dígitos $x_2, x_3, ..., x_7$

- la secuencia 01010, que se ubica en el centro

- la codificación binaria de los dígitos $x_8, ..., x_{13}$

- con 101 del lado derecho, que finaliza el código

Para códificar los dígitos $x_2, x_3, ..., x_7$ usaremos los códigos A y B, y un tercer código, que llamaremos C, para codificar los dígitos $x_8, ..., x_{13}$. El dígito x_1, no será codificado, pero condicionará el uso de los códigos A y B que se detallan a continuación.

x_1	x_2	x_3	x_4	x_5	x_6	x_7
0	A	A	A	A	A	A
1	A	A	B	A	B	B
2	A	A	B	B	A	B
3	A	A	B	B	B	A
4	A	B	A	A	B	B
5	A	B	B	A	A	B
6	A	B	B	B	A	A
7	A	B	A	B	A	B
8	A	B	A	B	B	A
9	A	B	B	A	B	A

dígito	código A	código B	código C
0	0001101	0100111	1110010
1	0011001	0110011	1100110
2	0010011	0011011	1101100
3	0111101	0100001	1000010
4	0100011	0011101	1011100
5	0110001	0111001	1001110
6	0101111	0000101	1010000
7	0111011	0010001	1000100
8	0110111	0001001	1001000
9	0001011	0010111	1110100

Consideremos la palabra código 77980044150555 del código EAN. Observemos que $x_1 = 7$, entonces, siguiendo la fila encabezado por el 7 en la primera tabla, x_2 se codifica con el código A, x_3 con el código B, x_4 con el código A y así sucesivamente. En la palabra código dada $x_2 = 7$ y de acuerdo a la segunda tabla, a 7 le corresponde el elemento 0111011 del código A, a $9 = x_3$ le corresponde el elemento 0010111 del código B. Sabiendo que los dígitos x_8 a x_{13} se codifican con el código C, se tiene que el EAN código de barras correspondiente a la palabra código dada es

101	extremo izquierdo
0111011 0010111 0110111 0100111 0100011 0011101	$x_2, x_3, ..., x_7$
01010	posición central
1100110 1001110 1110010 1001110 1001110 1001110	$x_8, ..., x_{13}$
101	extremo derecho

La representación en barras y espacios es

Figura 1.1.

1.4.4. El código ISBN

Cada libro publicado recientemente tiene asignado un número, una serie de 10 dígitos tomados del conjunto $\{0, 1, ..., 9, X\}$, llamado International Standard Book Number (ISBN). El código ISBN se asigna a las publicaciones monográficas (de sólo un

elemento físico y no publicaciones seriadas como lo son diarios y revistas) y, como es de esperarse, no son sólo 10 digitos escritos de forma aleatoria. Los 10 dígitos están divididos en 4 grupos, a veces separados por espacios o guiones, los cuales, su presencia o ausencia no mejoran ni perjudican el código. De izquierda a derecha, el primer dígito indica el idioma en que está escrito o el país de origen, los siguientes dos identifican al editor de la obra y es dado por la agencia nacional del ISBN, los siguientes seis dígitos son números que elige el editor y que identifican a una deretminada obra,y finalmente el último dígito está derterminado de la siguiente manera: supongamos que $x_1x_2...x_{10}$ es el ISBN de una obra. Estos dígitos deben satisfacer

$$\sum_{i=1}^{10} ix_i \equiv 0 \,(\text{mod } 11)$$

y así el último dígito, x_{10} queda determinado por

$$x_{10} = \sum_{i=1}^{9} ix_i (\text{mod } 11)$$

Si resulta que $x_{10} = 10$, entonces se escribe el símbolo X como dígito final, por ejemplo, $055010206X$. Existen formatos en donde la cantidad de dígitos los primeros tres grupos varían en su cantidad.
El código ISBN está diseñado para detectar

(a) un único error,

(b) dos errores producidos por la transposición de dos dígitos.

Para estos casos, veamos que no se cumple la condición que deben satisfacer los 10 dígitos. Supongamos que $x = x_1...x_{10}$ es el código ISBN de una obra y luego de enviar esta información se recibe la palabra $y = y_1...y_{10}$

(a) supongamos que ocurrió un error en el j-ésimo dígito y en vez de recibir x_j se recibió $x_j + a$ con $a \neq 0$, y que el resto de los dígitos de y son los mismos que los de x. Entonces

$$y = \sum_{i=1}^{10} iy_i = \left(\sum_{i=1}^{10} ix_i \right) + ja = ja \not\equiv 0 \,(\text{mod } 11)$$

debido a que j y a son no nulos.

(b) Supongamos que y es igual a x salvo los dígitos x_j y x_k que fueron transpuestos. Entonces, como la diferencia de dos elementos cualquiera de $\{1, .., 9\}$ no son mutiplos de 11 se tiene que

$$\begin{aligned} y = \sum_{i=1}^{10} iy_i &= \left(\sum_{i=1}^{10} ix_i \right) + (k-j)x_j + (j-k)x_k \\ &= (k-j)(x_j - x_k) \not\equiv 0 \,(\text{mod } 11) \end{aligned}$$

ya que $k \neq j$ y $x_j \neq x_k$.

Hasta diciembre de 2006 se ha usado el código ISBN con 10 dígitos y desde enero de 2007 se usa el código ISBN con 13 dígitos, que comienzan con la secuencia 978 y luego se sigue con los cuatro grupos citados anteriormente y cuando se agoten todos las combinaciones posibles con el comienzo 978 se comenzará con la secuencia 979. La forma de calcular el dígito número trece en este caso es, a partir de los doce primeros dígitos, de izquierda a derecha, multiplicar el primer dígito por 1, el segundo por 3, el cuarto por 1 y así sucesivamente hasta llegar al doceavo dígito. Luego se suman los resultados de las multiplicaciones anteriores y se calcula el resto r que resulta de dividir el resultado de la suma por 10. El último dígito será $10 - r$ si r es distinto de 0 o r en caso contrario.

2. Códigos lineales

En este capítulo estudiaremos a los códigos que están dotados de una estructura en particular, espacios vectoriales sobre cuerpos finitos. Esta estructura los hacen especialmente fáciles de caracterizar y muy útiles en la práctica debido a que la codificación y la decodificación corresponden a procedimientos simples y eficientes [4][5].

2.1. Códigos lineales

Sean $x = x_1x_2...x_n \in F_q^n$, $y = y_1y_2...y_n \in F_q^n$. La *suma de vectores* en F_q^n está definido por:

$$x + y = x_1 + y_1\ x_2 + y_2 \dots x_n + y_n.$$

y el *producto por un escalar* en F_q^n por:

$$\lambda x = \lambda x_1\ \lambda x_2 \dots \lambda x_n.$$

con $\lambda \in F_q$.
F_q^n, bajo la suma vectorial y el producto por un escalar definidos arriba, es un espacio vectorial.

Definición 2.1.1. (a) Un *código lineal* C de longitud n sobre F_q es un subespacio de F_q^n.

(b) La *dimensión* del código lineal C es la dimensión de C como espacio vectorial sobre F_q.

(c) Un código lineal C de longitud n y dimensión k es llamado un *código* $[n, k]$, y si además d es su distancia mínima, es llamado un *código* $[n, k, d]$.

Ejemplo 2.1.2. (a) $C_1 = \{00000, 01101, 10110, 11011\}$ es un código lineal y una base de él es $\{01101, 10110\}$. Por lo tanto C_1 es un código $[5, 2]$.

(b) $C_2 = \{(\lambda, \lambda, ..., \lambda) : \lambda \in F_q\}$ es un código lineal y es llamado código de repetición.

(c) $C_3 = \{0000, 1100, 2200, 0001, 0002, 1101, 1102, 2201, 2202\}$ es un código $[4, 2]$

(d) $C_4 = \{000, 011, 101, 110, 111\}$ no es un código lineal, porque por ejemplo $011 + 111 = 100$ no pertenece a C_4 y por lo tanto éste no es un espacio vectorial.

Definición 2.1.3. Sea x una palabra en F_q^n. El *peso (Hamming)* de x, denotado por $wt(x)$, está definido como el número de coordenadas no nulas en x, i.e.

$$wt(x) = d(x, 0)$$

con 0 la palabra nula en F_q^n.

Lema 2.1.4. *Si $x, y \in F_q^n$ entonces $d(x, y) = wt(x - y)$*

Demostración. Sea $x, y \in F_q^n$. Hagamos $x = x_1x_2...x_n$ y $y = y_1y_2...y_n$. Entonces

$$wt(x - y) = |\{i \,|\, x_i - y_i \neq 0,\ 1 \leq i \leq n\}| = |\{i \,|\, x_i \neq y_i,\ 1 \leq i \leq n\}| = d(x, y)$$

□

Definición 2.1.5. Sea C un código (no necesariamente lineal). El *peso (Hamming) mínimo* de C, denotado por $wt(C)$ es el menor de los pesos de las palabras códigos no nulas de C.

Teorema 2.1.6. *Sea C un código lineal sobre F_q. Entonces $d(C) = wt(C)$*

Demostración. Existen dos palablas códigos $x, y \in C$ tal que $d(x, y) = d(C)$. Por el lema anterior se tiene que

$$d(C) = wt(x - y) \geq wt(C)$$

Por otro lado, sea $z \in C$ la palabra código de peso mínimo

$$wt(C) = wt(z) = d(z, 0) \geq d(C)$$

Así como $d(C) \geq wt(C)$ y $wt(C) \geq d(C)$, se tiene que $wt(C) = d(C)$. □

Sabemos del álgebra lineal, que conocer una base de un espacio vectorial de dimensión finita nos permite describirlo totalmente. Un código lineal es un espacio vectorial de dimensión finita y para caracterizarlo debemos estudiar sus bases. Cada base suele estar ordenada en una matriz llamada matriz generadora del código y que presentamos luego del siguiente teorema.

Teorema 2.1.7. *Sea C un espacio vectorial sobre F_q. Si $\dim(C) = k$ entonces*

(a) C tiene q^k elementos.

(b) C tiene $\frac{1}{k!}\prod_{i=0}^{k-1}(q^k - q^i)$ bases distintas.

Demostración. (a) Sea $\{v_1 v_2 ... v_k\}$ una base de C, entonces

$$C = \{\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k \ : \lambda_1, \lambda_2, ..., \lambda_k \in F_q\}$$

y como $|F_q| = q$, hay q elecciones para cada λ_i con $i = 1...k$. Así C tiene q^k elementos.

(b) Sea $\{v_1 v_2 ... v_k\}$ una base de C.
Tenemos que $v_1 \neq 0$. Luego se tienen $q^k - 1$ elecciones para v_1.
Como los vectores de la base son linealmente independientes se observa que:
$v_2 \notin \langle v_1 \rangle = \{\lambda_1 v_1 \ : \lambda_1 \in F_q\}$ y como $|\langle v_1 \rangle| = q$ se tiene que hay $q^k - q$ elecciones para v_2.
$v_3 \notin \langle v_1, v_2 \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 \ : \lambda_1, \lambda_2 \in F_q\}$ y como $|\langle v_1, v_2 \rangle| = q^2$ se tiene que hay $q^k - q^2$ elecciones para v_3.
Así podemos concluir que, para $2 \leq i \leq k$ y $v_i \notin \langle v_1, ..., v_{i-1} \rangle$, hay $q^k - q^{i-1}$ elecciones para v_i.
Luego hay $\prod_{i=1}^{k}(q^k - q^{i-1})$ bases, o lo que es lo mismo, $\prod_{i=0}^{k-1}(q^k - q^i)$ bases y como el orden de los elementos de las bases no es relevante, se debe dividir esta última expresión por $k!$.
Con esto se concluye que hay $\frac{1}{k!}\prod_{i=1}^{k-1}(q^k - q^i)$ bases distintas.

□

Observación 2.1.8. Notar que si $|C| = q^k$, con C un espacio vectorial, entonces $\dim(C) = k$. Demostremos esto. Supongamos que $\dim(C) = s$, luego por teorema anterior $|C| = q^s$ y así $q^k = q^s$ con lo que resulta $k = s$ y por lo tanto $\dim(C) = k$.

Definición 2.1.9. Una *matriz generadora* de un código lineal C es una matriz G cuyas filas forman una base de C.

Ejemplo 2.1.10. Para cada código lineal C_i del Ejemplo (2.1.2) se tiene las siguientes matrices generadoras G_i para cada uno de ellos:

(a)
$$G_1 = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$
es una matriz generadora de $C_1 = \{00000, 01101, 10110, 11011\}$

(b) Suponiendo que la longitud de C_2 es 5, es decir, $C_2 = \{00000, 11111\}$
$$G_2 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

(c)
$$G_3 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
es una matriz generadora del código ternario $C_3 = \{0000, 1100, 2200, 0001, 0002, 1101, 1102, 2201, 2202\}$

El lector ya se habrá percatado que la matriz generadora de un código no es única, basta tener en cuenta el teorema (2.1.7) y la definición (2.1.1) para asegurar esto.

Ejemplo 2.1.11. Por teorema (2.1.7) el código lineal C_3 tiene 24 bases ya que:

$$\frac{1}{2}(3^2-1)(3^2-3)=24$$

Así por ejemplo $\{1100, 2202\}$, $\{1101, 2201\}$ y $\{0001, 1102\}$ son algunas otras bases de C_3 y por lo tanto

$$\begin{pmatrix}1&1&0&0\\2&2&0&2\end{pmatrix}\quad\begin{pmatrix}1&1&0&1\\2&2&0&1\end{pmatrix}\quad\begin{pmatrix}0&0&0&1\\1&1&0&2\end{pmatrix}$$

al igual que cualquier matriz formada a partir de las 24 bases calculadas, son también matrices generadoras de este código lineal.

2.2. Equivalencia de códigos lineales

Como ya dijimos, conocer la base de un código lineal, nos permite describirlo, describir sus palabras código.
Sea C un código $[n, k]$ y G una matriz generadora. Estamos interesados en que la matriz generadora del código tenga la forma

$$\left(\ I_k \mid A\ \right)$$

donde I_k es la matriz identidad de tamaño k y A es una matriz de $k \times (n-k)$. La matriz generadora en esta forma es llamada estándar, y ,dado un código, encontrarla será nuestro primer objetivo. Para ello va a ser de utilidad la siguiente definición que vincula cada código lineal con una base que forma una matriz generadora en forma estándar [5].

Definición 2.2.1. Dos códigos (n, M) sobre F_q son equivalentes si se puede obtener uno desde el otro por una combinación de las siguientes operaciones

(a) permutación de las posiciones de los símbolos de las palabras códigos;

(b) multiplicación de los símbolos que están en una posición fija por un escalar no nulo.

Ejemplo 2.2.2. Considerar los códigos lineales C_3 y C_5 con

$$C_3 = \{0000, 1100, 2200, 0001, 0002, 1101, 1102, 2201, 2202\}$$

$$C_5 = \{0000, 1001, 2002, 0100, 0200, 1101, 1201, 2102, 2202\}.$$

C_3 y C_5 son equivalentes, debido a que C_5 se obtiene de C_3 permutando el segundo con el cuarto símbolo.

Nota 2.2.3. Supongamos que $S = \{x_1, x_2, ..., x_k\}$ es una base del subespacio C de F_q^n. Es claro que podemos obtener otra base para C si:

(a) reemplazamos a x_i $(1 \leq i \leq k)$ por un múltiplo no nulo de el mismo, o

(b) reemplazamos a x_i por $x_i + ax_j$ para algun escalar a, con $1 \leq i,\ j \leq k$ y $j \neq i$.

Realicemos un esquema de demostración de esto. Sea $1 \leq i,\ j \leq k,\ j \neq i$ y, sin pérdida de generalidad, sea l tal que $i < l < j$.
Mostremos que $S' = \{x_1, ..., x_{i-1}, x_i + ax_j, ..., bx_l, ..., x_j, ..., x_k\}$ es una base de C.
Sea $L(S')$ el conjunto de todas las combinaciones lineales de los elementos de S', es decir

$$\begin{aligned}
L(S') &= \{a_1x_1 + ... + a_{i-1}x_{i-1} + a_i(x_i + ax_j) + a_{i+1}x_{i+1} + ... + a_l(bx_l) + ... + a_jx_j + \\
&\quad ... + a_kx_k \,|\, x_r \in S,\ a, b, a_r \in F_q,\ 1 \leq r \leq k\} \\
&= \{a_1x_1 + ... + a_ix_i + (a_ia)x_j + ... + (a_lb)x_l + ... + a_jx_j + ... + a_kx_k \,|\, x_r \in S, \\
&\quad a, b, a_r \in F_q,\ 1 \leq r \leq k\} \\
&= \{a_1x_1 + ... + a_ix_i + ... + (a_lb)x_l + ... + (a_j + a_ia)x_j + ... + a_kx_k \,|\, x_r \in S, \\
&\quad a, b, a_r \in F_q,\ 1 \leq r \leq k\} \\
&= \{a_1x_1 + ... + a_ix_i + ... + a'_lx_l + ... + a'_jx_j + ... + a_kx_k \,|\, x_r \in S, \\
&\quad a, b, a'_l, a'_ja_r \in F_q,\ 1 \leq r \leq k\}
\end{aligned}$$

La última igualdad nos dice que $L(S')$ es el conjunto de todas las combinaciones lineales de los elementos de S. Luego los elementos de S' generas a C. Queda como ejercicio para el lector demostrar que S' es un conjunto linealmente independiente. Así S' es base de C. Podemos afirmar entonces que dado un código lineal C y una base de él, realizando (a) y (b) obtenemos otra base de C.

Ejemplo 2.2.4. Consideremos que el código lineal $C_6 = \{000, 001, 100, 101\}$. C_6 tiene dimensión 2 y entonces, por Teorema (2.1.7), tiene tres bases:

$$\{001, 100\} \qquad \{001, 101\} \qquad \{100, 101\}$$

por lo tanto, las matrices generadoras son:

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

cada una formada ordenando los elementos de cada base en las filas de cada matriz. Observamos que aplicando las operaciones dadas en la nota anterior obtenemos las seis matrices generadoras de C_6 y que ninguna de ellas están en forma estándar.

Teorema 2.2.5. *Dos matrices de tamaño $k \times n$ generan códigos $[n, k]$ equivalentes sobre F_q si una matriz es obtenida de la otra por una sucesión de las operaciones del tipo:*

(R1) Permutación de filas.

(R2) Multiplicación de una fila por un escalar no nulo.

(R3) Adición del multiplo (no nulo) de una fila a otra.

(C1) Permutación de columnas.

(C2) Multiplicación de cualquier columna por un escalar no nulo.

Demostración. Con las operaciones (C1) y (C2) se se obtienen códigos lineales equivalentes por la definición (2.2.1). (R1), (R2) y (R3) se corresponden a lo afirmado en la nota (2.2.3). □

Observación 2.2.6. Notar que realizar las operaciones (R1), (R2) y (R3) implica obtener una base que genera el mismo código, mientras que realizar al menos una de las operaciones (C1) y (C2) implica obtener un código equivalente pero no el mismo.

Teorema 2.2.7. *Sea G una matriz generadora del código $[n, k]$. Entonces, aplicando las operaciones (R1), (R2), (R3), (C1) y (C2), G puede ser transformada en una matriz en forma estándard*

$$(\ I_k \ | \ A \ )$$

Donde I_k es la matriz identidad de tamaño k y A es una matriz de tamaño $k \times (n-k)$.

Demostración. Denotaremos con g_{ij} la (i, j)-ésima entrada de la matriz que resulta luego de la aplicación de cada operación y a $r_1, r_2, ..., r_k$ y $c_1, c_2, ..., c_n$ sus filas y columnas respectivamente.
Los tres pasos que se describen a continuación muestra cómo transformar la columna c_j en la forma deseada (con 1 en la posición j-ésima y $0s$ en las restantes) con $j = 1, 2, ..., k$, dejando sin cambios las $j - 1$ primeras columnas. Supongamos que G ha sido transformada en

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & g_{1j} & \cdots & g_{1n} \\ 0 & 1 & \cdots & 0 & g_{2j} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & g_{(j-1)j} & \cdots & g_{(j-1)n} \\ 0 & 0 & \cdots & 0 & g_{jj} & \cdots & g_{jn} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & g_{kj} & \cdots & g_{kn} \end{pmatrix}$$

Paso 1 Si $g_{jj} \neq 0$, continuamos con el Paso 2. Si $g_{jj} = 0$ y si para algún $i > j$ $g_{ij} \neq 0$ entonces intercambiamos r_j con r_i. Si $g_{jj} = 0$ y $g_{ij} = 0$ para todo $i > j$, entonces elegimos h tal que $g_{jh} \neq 0$ e intercambiamos c_j con c_h.

Paso 2 En el paso anterior obtuvimos $g_{jj} \neq 0$. Multiplicamos r_j por g_{jj}^{-1}.

Paso 3 En el paso anterior obtuvimos $g_{jj} = 1$. Para cada $i = 1, 2, ..., k$, con $i \neq j$, reemplazar r_i por $r_i - g_{ij}r_j$.

La columna c_j tienen ahora la forma deseada. Si repitimos estos pasos para $j = 1, 2, ..., k$ obtenemos una matriz generadora en su forma estándard. □

Observación 2.2.8. La matriz en forma estándar obtenida, debido a que se obtuvo a partir de las operaciones (R1), (R2), (R3), (C1) y (C2), puede ser que no genere el mismo código, pero si generará uno equivalente.
La matriz en forma estándard no siempre es única. Si aplicamos las operaciones (C1) y (C2) a las columnas de A (si se puede, dependiendo de la cantidad de columnas de A y del cuerpo sobre el que estamos trabajando) podemos obtener otra matriz en forma estándar.

Ejemplo 2.2.9. (a) Considerar el código C_3 y la matriz generadora del mismo

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 2 \end{pmatrix}$$

Obtengamos una matriz en forma estándar para un código equivalente a C_3.

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 2 \end{pmatrix} \xrightarrow{r_1 + r_2 \to r_2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$\xrightarrow{c_2 \to c_4} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{2c_2 \to c_2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(b) Encontremos una matriz generadora en forma estándar para un código lineal binario equivalente a un código lineal binario con matriz generadora

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

Procedamos a encontrar la matriz en forma estándar

$$\begin{pmatrix}1&0&1&0&1&0\\0&1&0&1&0&1\\1&1&0&1&1&0\\0&0&1&0&1&1\end{pmatrix} \xrightarrow[r_1+r_3\to r_3]{} \begin{pmatrix}1&0&1&0&1&0\\0&1&0&1&0&1\\0&1&1&1&0&0\\0&0&1&0&1&1\end{pmatrix}$$

$$\xrightarrow[r_2+r_3\to r_3]{} \begin{pmatrix}1&0&1&0&1&0\\0&1&0&1&0&1\\0&0&1&0&0&1\\0&0&1&0&1&1\end{pmatrix}$$

$$\xrightarrow[\substack{r_3+r_1\to r_1\\ r_3+r_4\to r_4}]{} \begin{pmatrix}1&0&0&0&1&1\\0&1&0&1&0&1\\0&0&1&0&0&1\\0&0&0&0&1&0\end{pmatrix}$$

$$\xrightarrow[c_5\to c_4]{} \begin{pmatrix}1&0&0&1&0&1\\0&1&0&0&1&1\\0&0&1&0&0&1\\0&0&0&1&0&0\end{pmatrix}$$

$$\xrightarrow[r_4+r_1\to r_1]{} \begin{pmatrix}1&0&0&0&0&1\\0&1&0&0&1&1\\0&0&1&0&0&1\\0&0&0&1&0&0\end{pmatrix}$$

2.3. Matriz de control de paridad, código dual

Dijimos que una forma de caracterizar a un código lineal es mediante el estudio de sus bases, lo que nos llevó encontrar ciertas matrices llamadas generadoras del código lineal. Las *matrices de control de paridad* también son de utilidad, sobre todo a la hora de decodificar el código [4][5]. Previo a introducir estas matrices, necesitaremos algunas definiciones.

Definición 2.3.1. Sea $x = x_1x_2...x_n,\ y = y_1y_2...y_n \in F_q^n$

(a) El *producto escalar* $x \cdot y$ (también conocido como *producto interno Euclideano*) está definido como

$$x \cdot y = x_1y_1 + x_2y_2 + ... + x_ny_n \in F_q.$$

(b) Dos elementos $x,\ y \in F_q^n$ se denominan *ortogonales* si $x \cdot y = 0$.

(c) Sea S un conjunto no vacío de F_q^n. El *complemento ortogonal* $S^\perp$ de S está definido como

$$S^\perp = \{x \in F_q^n : x \cdot y = 0 \textit{ para todo } y \in S\}$$

Definición 2.3.2. Sea C un código lineal en F_q^n. El *código dual* de C es $C^\perp$, i.e.

$$C^\perp = \{x \in F_q^n : x \cdot c = 0 \text{ para todo } c \in C\}$$

Ejemplo 2.3.3. teniendo en cuenta los códigos del Ejemplo (2.1.2)

(a) $C_1^\perp = \{00000, 10010, 01001, 00111, 11011, 10101, 01110, 11100\}$ es el código dual del código binario $C_1 = \{00000, 01101, 10110, 11011\}$.

(b) El código binario de repetición de longitud 4 es $C_2 = \{0000, 1111\}$. Se tiene que $C_2^\perp = \{0000, 1111, 1100, 1010, 1001, 0110, 0101, 0011\}$.

(c) $C_3^\perp = \{0000, 1200, 2100, 0010, 0020, 1210, 1220, 2110, 2120\}$es el código dual del código ternario $C_3 = \{0000, 1100, 2200, 0001, 0002, 1101, 1102, 2201, 2202\}$.

Probaremos que el código $C^\perp$ es un código lineal. Para esto será de utilidad el siguiente resultado, conocido del algebra lineal.

Lema 2.3.4. *Sea C un código $[n, k]$ con matriz generadora G. El elemento $x \in F_q^n$ pertenece a $C^\perp$ si y sólo si x es ortogonal a cada fila de G i.e. $x \in C^\perp \Leftrightarrow xG^\top = 0$, donde $G^\top$ es la matriz transpuesta de G.*

Demostración. Sean $r_1, r_2, ..., r_k$ las filas de G.
$\Rightarrow$) Como $\{r_1, r_2, ..., r_k\}$ es base de C, estas pertenecen a $r_1, r_2, ..., r_k \in C$ y como $x \in C^\perp$, x es ortogonal a cada r_i para $i = 1, 2, ..., k$.
$\Leftarrow$) Supongamos que $x \cdot r_i = 0$ para todo $i = 1, 2, ..., k$. Sea $c \in C$, entonces ocurre que

$$c = \lambda_1 r_1 + \lambda_2 r_2 + ... + \lambda_k r_k \text{ con } \lambda_i \in F_q \text{ para todo } i = 1, 2, ..., k$$

luego

$$x \cdot c = \textstyle\sum_{i_1}^{k} \lambda_i x \cdot r_i = 0.$$

de esto sigue la tésis. □

Teorema 2.3.5. *Sea C un código $[n, k]$ sobre F_q. El código dual $C^\perp$ de C es un código $[n, n-k]$.*

Demostración. Primero debemos probar que $C^\perp$ es un subespacio de F_q^n y concluir que es efectivamente un código lineal.
Sea $x, y \in C^\perp$, $\lambda_1, \lambda_2 \in F_q$ y $c \in C$. Entonces

$$(\lambda_1 x + \lambda_2 y)c = \lambda_1(xc) + \lambda_2(yc) = 0$$

y por lo tanto $\lambda_1 x + \lambda_2 y \in C^\perp$. Así $C^\perp$ es un código lineal.
Ahora determinaremos qué dimensión tiene $C^\perp$. Sin pérdida de generalidad supondremos que G, una matriz generadora de C, está en forma estándar

$$G = \begin{pmatrix} 1 & \cdots & 0 & a_{11} & \cdots & a_{1(n-k)} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & a_{k1} & \cdots & a_{k(n-k)} \end{pmatrix}$$

Sea $x = x_1 x_2 ... x_n \in F_q^n$. Si $x \in C^\perp$ entonces debe cumplir que $xG^\top = 0$

$$xG^\top = \begin{pmatrix} x_1 + \sum_{j=1}^{n-k} a_{1j} x_{k+j} & x_2 + \sum_{j=1}^{n-k} a_{2j} x_{k+j} & \cdots & x_k + \sum_{j=1}^{n-k} a_{kj} x_{k+j} \end{pmatrix}$$

Esto quiere decir que

$$C^\perp = \{x_1 x_2 ... x_n \in F_q^n : x_i + \sum_{j=1}^{n-k} a_{ij} x_{k+j} = 0, \quad i = 1, 2, ..., k\}$$

Hay q^{n-k} formas de armar $x_{k+1} ... x_n$ y una vez armado, existe un único vector $x_1 x_2 ... x_n$ que cumple con ser elemento de $C^\perp$. Así se tiene que $|C^\perp| = q^{n-k}$ y por la observación (2.1.8) $\dim(C^\perp) = n - k$. □

Teorema 2.3.6. *Sea C un código $[n, k]$, entonces*

(a) $\dim(C) + \dim(C^\perp) = n$.

(b) $(C^\perp)^\perp = C$.

Demostración. (a) Se deja para el lector.

(b) Sea $x \in C$, $y \in C^\perp$. Como $x \cdot y = 0$ se tiene que $x \in (C^\perp)^\perp$ y así, $C \subseteq (C^\perp)^\perp$. La $\dim(C^\perp) = n - k$ y por teorema anterior $\dim((C^\perp)^\perp) = n - (n - k) = k = \dim(C)$. De esto último sigue la tésis.

□

Definición 2.3.7. Sea C un código lineal.

(a) C es *auto ortogonal* si $C \subseteq C^\perp$.

(b) C es *auto dual* si $C = C^\perp$.

Ejemplo 2.3.8. (a) D_1, el código binario con matriz generadora A_1, y D_2, el código ternario con matriz generadora A_2, son códigos auto ortogonales.

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 2 \end{pmatrix}$$

En efecto, $D_1 = \{0000000, 0001111, 0110011, 1010101, 0111100, 1100110, 1011010, 1101001\}$ y

$D_1^{\perp} = \{0000000, 1110000, 0100101, 1000011, 0011001, 1001100, 0101110, 0010110, 1111111, 1101001, 1010101, 0110011, 1100110, 0111100, 1011010, 0001111\}$
donde se verifica que $D_1 \subset D_1^{\perp}$. Queda como ejercicio verificar que D_2 es un código auto ortogonal.

(b) D_3, el código binario con matriz generadora A_3, y D_4, el código ternario con matriz generadora A_4, son códigos auto duales.

$$A_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \quad A_4 = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

El código D_4 consiste en el conjunto $\{0000, 1011, 0121, 1102, 1220, 2110, 2022, 0212, 2201\}$. El lector puede verificar que el producto de dos palabras código cualquiera de D_4 da como resultado cero y que una base de $D_4^{\perp}$ es $\{1011, 1102\}$ y que ésta es a la vez base de D_4.

Teorema 2.3.9. *La dimensión de un código auto ortogonal de longitud n es menor o igual a $n/2$, y la dimensión de un código auto dual de longitud n es $n/2$.*

Demostración. Supongamos que C es un $[n, k]$ código auto ortogonal. Por teorema anterior $\dim(C) + \dim(C^{\perp}) = n$, y por la definición de código auto ortogonal, se tiene que

$$\begin{aligned} \dim(C) &\leq \dim(C^{\perp}) \\ k &\leq n - k \\ 2k &\leq n \\ k &\leq n/2 \end{aligned}$$

y así la dimensión de C es menor o igual a $n/2$.
Si C es auto dual, se cumple que $C \supseteq C^{\perp}$, así $k \geq n/2$, y además por ser auto ortogonal se tiene que $k \leq n/2$. Se concluye que la dimensión del código dual C es $n/2$ □

Definición 2.3.10. Una *matriz de control de paridad* H de C un código $[n, k]$ es una matriz generadora de $C^{\perp}$.

H es entonces una matriz de tamaño $(n-k) \times n$ cuyas filas son base de $C^{\perp}$ y además esta matriz satisface $GH^{\top} = 0$, donde 0 es la matriz nula. Luego podemos redefinir un código lineal C mediante su matriz de control de paridad como

$$C = \{x \in F_q^n | xH^{\top} = 0\}.$$

De esta forma cualquier código lineal está completamente determinado por una matriz de control de paridad.

Ejemplo 2.3.11. (a) Describamos el código lineal C binario con matriz de paridad

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Sea $x = x_1x_2x_3x_4x_5 \in F_2^5$. Calculemos $xH_1^\top$

$$xH_1^\top = \begin{pmatrix} x_1 + x_4 & x_2 + x_5 & x_3 + x_4 + x_5 \end{pmatrix} = \quad \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$$

Luego

$$\begin{aligned} C =& \{x_1x_2x_3x_4x_5 \in F_2^5 |\ xH_1^\top = 0\} \\ =& \{x_1x_2x_3x_4x_5 \in F_2^5 |\ x_1 + x_4 = 0,\ x_2 + x_5 = 0,\ x_3 + x_4 + x_5 = 0\} \\ =& \{x_1x_2x_3x_4x_5 \in F_2^5 |\ x_4 = x_1,\ x_5 = x_2,\ x_3 = x_1 + x_2\} \\ =& \{x_1x_2(x_1 + x_2)x_1x_2 \in F_2^5\} \\ =& \langle 10110, 01101 \rangle \end{aligned}$$

El lector podrá verificar que $C = C_1$ del ejemplo (2.1.2).

(b) Una matriz de control de paridad del código lineal C_3 del ejemplo (2.1.2) es

$$H_3 = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Luego, con $x_1x_2x_3x_4 \in F_3^4$, calculamos $xH_3^\top$

$$xH_3^\top = \begin{pmatrix} x_1 + 2x_2 & x_3 \end{pmatrix} = \quad \begin{pmatrix} 0 & 0 \end{pmatrix}$$

$$\begin{aligned} C_3 =& \{x_1x_2x_3x_4 \in F_3^4 |\ xH_1^\top = 0\} \\ =& \{x_1x_2x_3x_4 \in F_3^4 |\ x_1 + 2x_2 = 0,\ x_3 = 0\} \\ =& \{x_1x_2x_3x_4 \in F_3^4 |\ x_1 = x_2,\ x_3 = 0\} \\ =& \{x_1x_10x_4 \in F_3^4\} \\ =& \langle 1100, 0001 \rangle \end{aligned}$$

Las ecuaciones igualadas a cero en el ejemplo anterior tienen el nombre de *ecuaciones de control de paridad.*

Lema 2.3.12. *Sea C un código $[n, k]$ sobre F_q, con G una matriz generadora. Entonces*

(a) Sea $x \in F_q^n$. x pertenece a $C^\perp$ si y sólo si x es ortogonal a cada fila de G.

(b) Dada una matriz H de tamaño $(n-k)\times n$ entonces, H es una matriz de control de paridad de C si y sólo si las filas de H son linearmente independientes y $HG^\top = 0$

Demostración. El inciso (a) queda como ejercicio. Demostremos el inciso (b).
$\Rightarrow$) Como H es la matriz de control de paridad de C sus filas son linealmente independientes, además éstas son elementos de $C^{\perp}$ y por lo tanto se cumple que $HG^{\top} = 0$.
$\Leftarrow$) Como se cumple que $HG^{\top} = 0$ se tiene por inciso (a) que las filas de H están contenidas en $C^{\perp}$ y éstas por hipótesis son linealmente independientes por lo que generan un subespacio de dimensión $(n-k)$. Con esto concluimos que las filas de H generan a $C^{\perp}$ y por lo tanto es una matriz de control de paridad de C. □

Teorema 2.3.13. *Si $G = (I_k|A)$ es una matriz generadora en forma estándar de C un código $[n,k]$, entonces una matriz de control de paridad de C es $H = (-A^{\top}|I_{n-k})$.*

Demostración. Supongamos

$$G = \begin{pmatrix} 1 & & 0 & a_{11} & \cdots & a_{1(n-k)} \\ & \ddots & & \vdots & & \vdots \\ 0 & & 1 & a_{k1} & \cdots & a_{k(n-k)} \end{pmatrix}$$

y sea

$$H = \begin{pmatrix} -a_{11} & \cdots & -a_{k1} & 1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ -a_{1(n-k)} & \cdots & -a_{k(n-k)} & 0 & & 1 \end{pmatrix}$$

Como las filas de H son independientes y cumple con el tamaño que debe tener para ser una matriz de control de paridad, para demostrar este teorema basta probar que $GH^{\top} = 0$.
La i-ésima fila de G es

$$0 \ldots 1 \ldots 0 \; a_{i1} \; a_{i2} \ldots a_{i(n-k)} \qquad \text{con 1 en la cordenada } i$$

y la j-ésima fila de H es

$$-a_{1j} \; -a_{2j} \ldots \; -a_{(n-k)j} \; 0 \ldots 1 \ldots 0 \qquad \text{con 1 en la cordenada } k+j$$

Luego el producto escalar entre la i-ésima fila de G y la j-ésima fila de H resulta

$$0 + \ldots + 0 + (-a_{ij}) + 0 + \ldots + 0 + a_{ij} + 0 + \ldots + 0 = 0$$

de esto sigue la tésis. □

Una matriz de control de paridad H de C, un código $[n,k]$, está en forma estándard si $H = (B|I_{n-k})$.
La prueba del teorema anterior nos permite encontrar la matriz generadora de C a partir de su matriz de control de paridad en forma estándar, y ésta es $G = (I_k|-B^{\top})$.

Dada la matriz de control de paridad, podemos encontrar una en forma estándar, de la misma forma que encontramos la forma estándar de la matriz generadora del código, y una vez encontrada, ésta puede no pertenecer al código original, pero si a uno equivalente.

Ejemplo 2.3.14. Consideremos el código $[10, 8]$ sobre F_{11} con matriz generadora

$$G = \begin{pmatrix} & & 2 & 8 \\ & & 3 & 7 \\ & & 4 & 6 \\ & I_8 & 5 & 5 \\ & & 6 & 4 \\ & & 7 & 3 \\ & & 8 & 2 \\ & & 9 & 1 \end{pmatrix} \text{ donde } A = \begin{pmatrix} 2 & 8 \\ 3 & 7 \\ 4 & 6 \\ 5 & 5 \\ 6 & 4 \\ 7 & 3 \\ 8 & 2 \\ 9 & 1 \end{pmatrix}$$

Luego buscamos $-A^\top$

$$-A^\top = \begin{pmatrix} -2 & -3 & -4 & -5 & -6 & -7 & -8 & -9 \\ -8 & -7 & -6 & -5 & -4 & -3 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 \\ 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{pmatrix}$$

y finalmente

$$H = \begin{pmatrix} 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 0 & 1 \end{pmatrix}$$

El siguiente teorema relaciona la matriz de control de paridad con la distancia mínima y es de utilidad cuando esta última es pequeña.

Teorema 2.3.15. *Sea C un código lineal y H una matriz de control de paridad de C. Entonces*

(a) C tiene distancia mayor o igual a d si y sólo si cualquier conjunto de $d-1$ columnas de H son linealmente independientes,

(b) C tiene distancia menor o igual que d si y sólo si H tiene d columnas que son linealmente dependientes.

Demostración. (a) Sea $x = x_1x_2...x_n \in C$ una palabra código de peso $e > 0$. Supongamos que las coordenadas no nulas están en las posiciones $i_1, ..., i_e$, i.e. $x_j = 0$ si y sólo si $j \notin \{i_1, ..., i_e\}$. Sea c_l $(1 \leq l \leq n)$ la l-ésima columna de H. Por lema (2.3.12) $x \in C$ si y sólo si,

$$xH^\top = 0 = x_{i_1}c_{i_1}^\top + ... + x_{i_e}c_{i_e}^\top$$

lo cual es cierto si estas e columnas de H son linealmente dependientes. Ahora suponer que la distancia de C es mayor o igual a d es equivalente a decir que C no tiene palabras código con peso menor o igual a $d-1$ que es equivalente a decir que cualquier conjunto de $d-1$ columnas de H son linealmente independientes, ya que si fueran linealmente dependientes habría una palabra código de peso $d-1$ en C.

(b) De forma análoga, afirmar que C tiene distancia menor o igual a d es equivalente a decir que C contiene una palabra no nula de peso menor o igual a d, que es equivalente a decir que H tiene d o menos columnas linealmente dependientes (y por lo tanto d).

□

Corolario 2.3.16. *Sea C un un código lineal y H una matriz de control de paridad de C. Las siguientes afirmaciones son equivalentes:*

(a) C tiene distancia d

(b) cualquier conjunto de $d-1$ columnas de H son linealmente independientes y H tiene d columnas que son linealmente dependientes.

Sea C el código lineal binario con matriz de control de paridad

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

Observamos que las columnas 1, 2 y 6 son linealmente dependientes. Por otro lado, cualquier conjunto formado por dos conjuntos de H es linealmente independiente. Así la distancia mínima de C es $d = 3$.

2.4. Codificación y decodificación

2.4.1. Codificación con un código lineal

Consideremos a C, un código $[n, k]$ sobre F_q con matriz generadora G. Como $|C| = q^k$, podemos transmitir q^k mensajes distintos. Para este código identificaremos a los mensajes con k-uplas, i.e. si x es un mensaje, entonces $x \in F_q^k$.
Sea $x = x_1x_2...x_k \in F_q^k$ y f definida por

$$f : F_q^k \to F_q^n |\ f(x) = xG.$$

donde G es la matriz generadora de C, un $[n, k]$ código. Si $r_1,\ r_2,\ ...,\ r_k$ son las filas de G, entonces

$$xG = \sum_{i=1}^{k} x_i r_i$$

es una palabra código ya que es una combinación lineal de las filas de G.

Codificar un mensaje $x \in F_q^k$ con el codigo lineal C significa aplicar a x la función f.

Ejemplo 2.4.1. (a) Consideremos a

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

una matriz generadora en forma estándar de un código lineal C equivalente a C_3. Si por ejemplo se desea enviar el mensaje $x = 21$ entonces la codificación del mismo es

$$xG = 2(1001) + 0100 = 2102$$

La siguiente tabla muestra la palabra código correspondiente a los mensajes $00, 01, 01, 11, 02, 20, 22, 12, 21$

Mensaje	Palabra código
00	0000
10	1001
01	0100
11	1101
20	2002
02	0200
22	2202
12	1201
21	2102

(b) Considerar a C un código binario $[7, 4]$ con matriz generadora

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

Entonces la correspondencia entre mensajes y palabras código está dada por las siguientes tablas

Mensaje	Palabra código
0000	0000000
1000	1000011
0100	0100101
0010	0010110
0001	0001111
1100	1100110
1010	1010101
1001	1001100

Mensaje	Palabra código
0110	0110011
0101	0101010
0011	0011001
1110	1110000
1101	1101001
1011	1011010
0111	0111100
1111	1111111

Observación 2.4.2. Consideremos un $[n, k]$ código C con matriz generadora G en forma estándar, i.e. $G = (I_k|A)$, donde $A = (a_{ij})$ es una matriz de $k \times (n-k)$. Sea $x = x_1x_2...x_k \in F_q^k$ tal que $f(x) = c$, con $c = c_1c_2...c_n$. Entonces

$$xG = (x_1x_2...x_k) \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & A \\ & & 1 & \end{array} \right) = c_1c_2...c_n$$

donde $c_i = x_i,\ 1 \leq i \leq k$, i.e. consiste en el mensaje mismo y

$$c_{k+i} = \sum_{j=1}^{k} x_j a_{ji}, \quad 1 \leq i \leq n-k$$

estos últimos $n-k$ dígitos se llaman *dígitos de control* y representa la *redundancia* que ha sido agregada para la protección del mensaje contra el ruido. Redundancia son dígitos que se agregan al mensaje; cuando codificamos le asignamos al mensaje una palabra código que tiene una longitud mayor a la del mensaje, independientemente la matriz generadora que usamos. Así cuando codificamos un mensaje le agregamos redundancia.
En el ejemplo anterior el lector podrá comprobar lo dicho en esta observación.

2.4.2. Decodificación de un código lineal

Supongamos que la palabra código $c = c_1c_2...c_n$ es enviada a través de un canal simétrico y que la palabra recibida es $y = y_1y_2...y_n$. Definimos el *error* o *vector error* e como

$$e = y - c = e_1e_2...e_n.$$

El decodificador debe decidir a partir de y cuál fue la palabra código c transmitida o, de forma equivalente, cuál fue el error e que ocurrió. Un código tiene un uso práctico si su decodificacion es eficiente. Aunque es usual que cada código tenga un esquema de decodificación propio, mostraremos aquí un esquema de decodificación simple pero elegante para códigos lineales, aporte realizado por el matemático estadounidense David Slepian, quien usa el hecho de que un código lineal es un subgrupo del grupo aditivo F_q^n. Esto se debe a que F_q^n es un espacio vectorial y la operación suma que le da esta estructura, cumple con las condiciones para que F_q^n sea un grupo abeliano. También mostraremos la modificación de este esquema para mejorar su rendimiento cuando la longitud y tamaño del código sean grande [5].

Definición 2.4.3. Sea C un código $[n, k]$ sobre F_q y $a \in F_q^n$. Entonces el conjunto $a + C$ definido por

$$a + C = \{a + x \,|\, x \in C\}$$

es llamado *coclase* de C.

Teorema 2.4.4. *Sea C un código $[n, k]$ sobre F_q. Entonces*

(a) todo elemento de F_q^n está contenido en alguna coclase de C,

(b) para todo $a \in F_q^n$, $|a + C| = |C| = q^k$,

(c) para todo $a, b \in F_q^n$, $a \in b + C$ si y sólo si $a + C = b + C$,

(d) dos coclases o son disjuntas o iguales,

(e) existen q^{n-k} coclases diferentes de C,

(f) para todo $a, b \in F_q^n$, $a - b \in C$ si y sólo si a y b están en la misma coclase.

Demostración. (a) Sea $a \in F_q^n$, entonces $a = a + 0 \in a + C$, con 0 el vector nulo y perteneciente a C por ser éste un espacio vectorial.

(b) Puesto que el vector nulo pertenece a C, $a + C$ tiene a lo sumo $|C| = q^k$ elementos.
Ahora dos elementos $a + c$ y $a + c'$ de $a + C$ son iguales si y solo si $c = c'$. Esto nos conduce a afirmar que $|a + C| = q^k = |C|$.

(c) Como $a \in b + C$, existe $x \in C$ tal que $a = b + x$. Sea $d \in a + C$, existe $y \in C$ tal que $d = a + y$. Así

$$d = a + y = (b + x) + y = b + (x + y) \in b + C,$$

y como $x, y \in C$, $x + y \in C$ y se tiene que $d \in b + C$. Luego $a + C \subseteq b + C$.
Por otro lado, sea $f \in b + C$, existe $z \in C$ tal que $f = b + z$, entonces

$$f = b + z = (a - x) + z = a + (z - x) \in a + C$$

ya que $x, -x, z \in C$. Luego $b + C \subseteq a + C$ y así $b + C = a + C$.
Se deja como ejercicio para el lector demostrar la implicación recíproca.

(d) Para demostrar este ítem, partiremos del supuesto de que dos coclases no son disjuntas y demostraremos, a partir de este supuesto, que si estas coclases no son disjuntas, entonces son iguales. Supongamos que las coclases $a + C$ y $b + C$ de C no son disjuntas, esto significa que existe un elemento $v \in F_q^n$ tal que $x \in (a + C) \cap (b + C)$. Luego existen $y, z \in C$ tal que

$$x = a + y = b + z$$

Así $b = a + (y - z) \in a + C$ y por el item anterior $a + C = b + C$.

(e) De (a),(b) y (d) sigue la tésis.

(f) $\Rightarrow$) Como $a - b = c \in C$ se tiene que $a = b + c$ i.e. $a \in b + C$. Por (a), $a \in a + C$ y por (c) $a + C = b + C$. Por lo tanto, a y b pertenecen a la misma coclase.
$\Leftarrow$) Supongamos que a y b están en la misma coclase $x + C$. Entonces existen $y, z \in C$ tales que $a = x + y$ y $b = x + z$. Así $a - b = y - z \in C$.

□

Definición 2.4.5. Una palabra con el menor peso en una coclase es llamada *representante* de la coclase.

Ejemplo 2.4.6. (a) Sea $C = \{00000, 01101, 10110, 11011\}$ un código $[5, 2]$. Construyamos la coclase de 10000.

$$10000 + 00000 = 10000 \qquad 10000 + 01101 = 11101$$

$$10000 + 10110 = 00110 \qquad 10000 + 11011 = 01011$$

Entonces $10000 + C = \{10000, 11101, 00110, 01011\}$. Las coclases de C son:

$$\begin{aligned}
00000 + C &= \{00000, 01101, 10110, 11011\} \\
10000 + C &= \{10000, 11101, 00110, 01011\} \\
01000 + C &= \{01000, 00101, 11110, 10011\} \\
00100 + C &= \{00100, 01001, 10010, 11111\} \\
00010 + C &= \{00010, 01111, 10100, 11001\} \\
00001 + C &= \{00001, 01100, 10111, 11010\} \\
11000 + C &= \{11000, 10101, 01110, 00011\} \\
10001 + C &= \{10001, 11100, 00111, 01010\}
\end{aligned}$$

(b) Sea $D = \{0000, 1011, 0101, 1110\}$ un código $[4, 2]$, entonces sus coclases son:

$$\begin{aligned}
0000 + D &= \{0000, 1011, 0101, 1110\} \\
1000 + D &= \{1000, 0011, 1101, 0110\} \\
0100 + D &= \{0100, 1111, 0001, 1010\} \\
0010 + D &= \{0010, 1001, 0111, 1100\}
\end{aligned}$$

Observación 2.4.7.

1. Notemos que $01111 + C = \{01111, 00010, 10100, 11001\}$ y resulta que $01111 + C = 00010 + C$, también las coclases $10100 + C$ y $11001 + C$ son iguales a la coclase $00010 + C$ y así se verifica lo afirmado en el teorema anterior.
2. Una coclase puede tener más de un representante. Por ejemplo, la coclase $11000 + C$ tiene a 11000 y 00011 como representantes; la coclase $10001 + C$ tiene a 10001 y a 01010 como representantes y la coclase $0100 + D$ tiene como representantes a 0100 y a 0001.

Nota 2.4.8. Usualmente los incisos (a), (b) y (d) se presentan en un teorema aparte que se llama Teorema de Lagrange. Si C es un código $[n, k]$, F_q^n queda particionado y es la unión (disjunta) de las coclases de C, i.e.

$$F_q^n = (0 + C) \cup (a_1 + C) \cup ... \cup (a_s + C)$$

donde $s = q^{n-k} - 1$ y $a_1, a_2, ..., a_s$ son los representantes de cada coclase.

Usaremos las coclases de un código lineal para describir un nuevo método de decodificación.
Una *tabla estándar* (de Slepian) de C, un código $[n, k]$ sobre F_q^n, es una tabla de tamaño $q^{n-k} \times q^k$ que contiene todos los elementos de F_q^n, en el que la primera fila consiste en el código C, con $0 \in C$ en el extremo izquierdo, y las demás filas se ubican los elementos de las coclases $a_i + C$ con a_i, el representante de cada coclase, en el extremo izquierdo de cada fila.
Podemos armar esta tabla siguiendo los pasos a continuación:

Paso 1 Ubicar las palabras código de C en la primera fila, con $0 \in C$ en el extremo izquierdo.

Paso 2 Elegir cualquier elemento $a_1 \in F_q^n$, que no esté en la primera fila y que tenga el menor peso que el resto de las palabras, que cumpla

$$w(a_1) \leq \left\lfloor \frac{d-1}{2} \right\rfloor$$

Ubicar los elementos de la coclase $a_1 + C$ poniendo a_1 debajo de 0 y $a_1 + x$ debajo de cada x, con $x \in C$.

Paso 3 Elegir cualquier elemento $a_2 \in F_q^n$, que no esté en la primera ni en la segunda fila y que tenga el menor peso que el resto de las palabras,que cumpla

$$w(a_2) \leq \left\lfloor \frac{d-1}{2} \right\rfloor.$$

Ubicamos los elementos de $a_2 + C$ siguiendo el paso 2.

Paso 4 Continuar de esta forma hasta ubicar los elementos de todas las coclases.

Ejemplo 2.4.9. Consideremos los códigos lineales C y D del ejemplo (2.4.6)

(a) El primer paso para construir la tabla es ubicar las palabras código de C en la primera fila, con $0 \in C$ en el extremo izquierdo. Luego elijamos a 10000 que cumple que $10000 \leq \left\lfloor \frac{3-1}{2} \right\rfloor$ y hacemos

00000	01101	10110	11011
10000	10000 + 01101	10000 + 10110	10000 + 11011

Luego las primeras dos filas de la tabla estándar de C son

00000	01101	10110	11011
10000	11101	00110	01011

La tabla estándar para C es

código →	00000	01101	10110	11011
	10000	11101	00110	01011
	01000	00101	11110	10011
	00100	01001	10010	11111
	00010	01111	10100	11001
	00001	01100	10111	11010
	↑			
	representantes			

(b) La tabla estándar para D es

0000	1011	0101	1110
1000	0011	1101	0110
0100	1111	0001	1010
0010	1001	0111	1100

Denotando por a_{ij} el elemento de la tabla estándar ubicado en la fila i y la columna j, $a_{ij} = a_{i1} + a_{1j}$, con a_{i1} el representante de la coclase ubicada en la fila i de la tabla y a_{1j} la palabra código ubicada en la primera fila en la j-ésima posición. Así construimos la tabla, por lo que es correcto decir que $a_{1j} = a_{ij} - a_{i1}$.
La decodificación por medio de esta tabla es la siguiente:
Suponiendo que la palabra y es recibida, identificamos su posición en la tabla estándar. El decodificador decide que el error cometido en la transmisión es el representante e de la coclase a la que pertenece y e y es decodificado como $c = y - e$. En pocas palabras y se decodifica como la palabra código que se encuentra en lo alto de la columna donde se ubica y en la tabla estándar.
Como dijimos, el error cometido es el representante de cada coclase que por definición es el elemento de la coclase con menor peso. Esta elección nos permite decir que esta forma de decodificar no es otra cosa que la decodificación de la distancia mínima.
Consideremos el código D del ejemplo anterior. Supongamos que la palabra 1010 es recibida. Mirando la tabla y aplicando la decodificación por medio de la misma, sabemos que el error producido es 0100 y que 1010 se decodifica como 1110. Recordamos que las coclases pueden tener más de un posible representante, y que a este lo elegimos de forma arbitraria para construir la tabla. Así en este caso hay dos posibles representantes, y elegido fue 0100. Sin embargo si se elegia como representante a 0001, 1010 se decodificaba como una palabra código distinta a 1110.
El lector se preguntará ¿qué se hace en este caso? En el capítulo 1 describimos dos tipos de decodificación de la distancia mínima, la completa y la incompleta. Si elegimos realizar la decodificación completa, entonces simplemente elegimos arbitrariamente un representante de los posibles y así el error producido. Si elegimos realizar la decodificación incompleta, en caso que la palabra recibida pertenezca a una coclase con más de un representante posible, pedimos una retransmisión.

Nota 2.4.10. En la práctica, la decodificación por medio de la tabla estándar, es muy lenta para códigos de longitud y tamaño muy grandes (y también costoso en

términos de requerimientos de almacenamiento). Por eso, describiremos una forma decodificar con la tabla estándar llamada *decodificación por síndrome* y que describimos a continuación.

2.4.3. Decodificación por síndrome

Definición 2.4.11. Sea C un código $[n, k, d]$ sobre F_q y sea H una matriz de control de paridad de C. Para todo $y \in F_q^n$, el *síndrome* de y es $S(y) = yH^\top$.

Nota 2.4.12. Sabemos que para un código C no hay una única matriz de control de paridad, por lo tanto sería más conveniente denotar al síndrome de y por $S_H(y)$, para hacer notar la dependencia de la matriz de control de paridad. Sin embargo para simplificar notación, no usaremos el subíndice H asumiendo que no hay lugar a confusión.

Teorema 2.4.13. *Sea C un código $[n, k, d]$ sobre F_q y sea H una matriz de control de paridad de C. Si $x, y \in F_q^n$ entonces*

(a) $S(x + y) = S(x) + S(y)$,

(b) $S(x) = 0$ *si y sólo si* $x \in C$,

(c) x *e* y *están en una misma coclase de* C *si y sólo si* $S(x) = S(y)$.

Demostración. Los incisos (a) y (b) quedan como ejercicio para el lector.
Sean $x, y \in F_q^n$ que están en una misma coclase de C, esto es

$$\begin{aligned} & x + C = y + C \\ \Leftrightarrow & x - y \in C \\ \Leftrightarrow & (x - y)H^\top = 0 \\ \Leftrightarrow & xH^\top = yH^\top \\ \Leftrightarrow & S(x) = S(y). \end{aligned}$$

□

Corolario 2.4.14. *Existe una correspondencia uno a uno entre coclases y síndromes.*

Construiremos una nueva tabla a la que también llamaremos tabla estándar, asumiendo la decodificación completa, siguiendo los siguientes pasos:

Paso 1 Listar todas las coclases y elegir de cada una un representante.

Paso 2 Encontrar una matriz de control de paridad H para el código lineal y calcular $S(x) = xH^\top$, siendo x los representantes de cada coclase.

Paso 3 Construir una columna con los representantes de cada coclase, que cumplan con tener peso menor o igual que $\lfloor \frac{d-1}{2} \rfloor$ comenzando con $0 \in C$ y otra con los síndromes correspondientes a cada representante.

En el caso de asumir la decodificación incompleta, si una coclase tiene más de un posible representante, no escribiremos un representante en la tabla, sino el símbolo '$*$' en su lugar. Además si escribiremos el síndrome que le corresponde a los elementos de esa coclase.
El siguiente ejemplo ilustra la construcción de las tablas.

Ejemplo 2.4.15. Sea $D = \{0000, 1011, 0101, 1110\}$ el código lineal del ejemplo (2.4.6). Si consideramos a

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

como una matriz de control de paridad de D, entonces las tablas correspondientes asumiendo la decodificación completa y la incompleta son:

Representante	Síndrome
0000	00
1000	11
0100	01
0010	10

Representante	Síndrome
0000	00
1000	11
$*$	01
0010	10

El procedimiento de decodificación, asumiendo la decodificación completa, utilizando estas tablas es el siguiente

Paso 1 Si se recibió la palabra y, calculamos $S(y)$.

Paso 2 Encontrar el representante e, tal que $S(y) = S(e)$.

Paso 3 Decodificar y como $c = y - e$

Si asumimos la decodificación incompleta, seguimos los pasos anteriores a menos que el representante en el paso 2 no sea encontrado, i.e. nos encontremos con un '$*$' en la busqueda del representante. en este caso solicitaremos la retransmisión del mensaje.

Ejemplo 2.4.16. Sea $C = \{00000, 01101, 10110, 11011\}$ un código $[5, 2, 3]$. Consideremos la siguiente matriz de control de paridad de H

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

Construyamos la tabla estándar usando el síndrome

Representante	Síndrome
00000	000
10000	110
01000	101
00100	100
00010	010
00001	001

Podemos observar que en la columna de los representantes no se encuentran los representantes de las coclases $11000 + C$ y $10001 + C$. Esto se debe a que el ninguno de los representantes de estas coclase cumple con tener peso menor o igual a $\lfloor \frac{d-1}{2} \rfloor$, que en este caso es 1.

Decodificar las siguientes palabras
a) 10100 b) 11111 c) 10111 d) 00101 e) 01111
Para poder decodificar estas palabras, procedemos a calcular el síndrome de cada palabra y utilizar la tabla para encontrar el error cometido y corregirlo.

(a) $10100H^\top = 010$
De acuerdo a la tabla el representante que tiene síndrome 010 es 00010. Por lo tanto 10100 se decodifica como $10100 - 00010 = 10110$.

(b) $11111H^\top = 100$
La tabla nos dice que el representante con síndrome 100 es 00100. Por lo tanto 11111 se decodifica como $11111 - 00100$, i.e. como 11011.

(c) $10111H^\top = 001$
Por lo tanto 10111 se decodifica como 10110.

(d) $11000H^\top = 011$
El síndrome encontrado no se encuentra en la tabla, por lo que es necesario pedir una retransmisión del mensaje.

(e) $01111H^\top = 010$
Por lo tanto 10100 se decodifica como 01101.

Teorema 2.4.17. *Sea C un código lineal. La decodificación por síndrome es equivalente a la decodificación de la distancia mínima.*

Demostración. Supongamos que se recibió la palabra y. Sea e el representante de la coclase $y + C = e + C$. Así se tiene que $y = e + c$, con $c \in C$. La decodificación por síndrome dice que y se decodifica como $y - e$ i.e. como c.
Luego como $d(y, c) = wt(y - c) = wt(e)$, se tiene que al minimizar el peso de e, minimizamos la distancia de y a una palabra código c, con $c \in C$ [7]. □

Nota 2.4.18. A continuación se describen las ventajas y desventajas de utilizar centrar la atención a los códigos lineales

- *Ventaja 1* Para un codigo general (n, M), para encontrar su distancia mínima debemos hacer $\binom{M}{2} = \frac{M(M-1)}{2}$ comparaciones. En cambio si dicho código fuera lineal, solo debemos encontrar la palabra con peso mínimo, i.e. examinar las $M - 1$ palabras no nulas.
- *Ventaja 2* Para describir un código no lineal, debemos listar todas sus palabras código. En cambio, para describir un código únicamente necesitamos conocer una de sus bases.
- *Ventaja 3* Existen muy buenos procedimientos para codificar y decodificar códigos lineales.
- *Desventaja 1* Un código lineal debe estar definido sobre un cuerpo finito, para que este tenga la estructura de espacio vectorial. Si F_q es el cuerpo sobre el que se construye el código, entonces q es la potencia de un número primo, para que F_q tenga la estructura de cuerpo. Esto representa una limitación en cuanto a la elección del alfabeto.

3. Cotas en la teoría de códigos

3.1. El principal problema de la teoría de códigos

Un buen código (n, M, d) tiene a n pequeño para una rápida transmisión, M grande para una amplia variedad de mensajes y d grande para corregir la mayor cantidad de errores. El *principal problema de la teoría de códigos* es optimizar uno de los parámetros n, M o d fijando los dos parámetros restantes. La versión original de este problema es encontrar el código con más palabras código dados n y d [5].

Definición 3.1.1. Sea A un alfabeto con q elementos ($q > 1$) y los valores fijos n y d. $A_q(n, d)$ denota el valor más grande posible de M tal que existe el código (n, M, d) sobre A i.e.

$$A_q(n, d) = \text{máx}\{M| \text{ existe un } (n, M, d) \text{ código sobre } A\}$$

Notemos que $A_q(n, d)$ depende del tamaño del alfabeto, de n y d. Cualquier código (n, M, d) que cumple con $M = A_q(n, d)$ es llamado *código óptimo* [4].
De forma particular, para códigos lineales tenemos la siguiente definición.

Definición 3.1.2. Sea q la potencia de un número primo y valores fijos n y d. $B_q(n, d)$ denota el valor más grande posible de q^k para el cual existe el código $[n, k, d]$ sobre F_q i.e.

$$B_q(n, d) = \text{máx}\{q^k| \text{ existe un } [n, k, d] \text{ código sobre } F_q\}$$

Determinar el valor $A_q(n, d)$ (y de $B_q(n, d)$) es muy difícil, sin embargo, es posible acotar este valor. En este capítulo presentamos algunas cotas inferiores y superiores para $A_q(n, d)$ (y para $B_q(n, d)$) y describiremos algunos códigos particulares llamados perfectos.

Teorema 3.1.3. *Sea $q \geq 2$ la potencia de un número primo. Entonces*

(a) $B_q(n, d) \leq A_q(n, d) \leq q^n$ *para todo* $1 \leq d \leq n$.

(b) $B_q(n, 1) = A_q(n, 1) = q^n$.

(c) $B_q(n, n) = A_q(n, n) = q$.

Demostración. (a) $B_q(n,d) \leq A_q(n,d)$ se cumple puesto que una vez encontrado $B_q(n,d)$ basta agregarle a ese código una palabra de F_q^n de peso mayor o igual a d. También se cumple de forma trivial $A_q(n,d) \leq q^n$, ya que q^n es la cantidad total de palabras de F_q^n.

(b) Notemos que F_q^n es un código $[n,n,1]$ por lo tanto $B_q(n,1) = q^n$. Luego por (a) se tiene la tésis.

(c) Sea C un código (n,M,n) sobre F_q. Ya que las palabras código son de longitud n y la distancia entre cualquiera de ellas es mayor o igual que n, se sigue que la distancia entre cualquier par de palabras código es exactamente n i.e. las palabras difieren en todas sus coordenadas. Así debe ocurrir que las coordenadas de una palabra código deben ser iguales, y distintas a las coordenadas de otra palabra código, en otras palabras, C es el código de repetición de longitud n. Esto implica que la cantidad de palabras es q y así $B_q(n,n) \leq A_q(n,n) \leq q$ y como el código de repetición es lineal se tiene que $B_q(n,n) = q = A_q(n,n)$.

□

En el caso de los códigos sobre F_2 hay resultados elementales para determinar $A_q(n,d)$ y $B_q(n,d)$. Para discutirlos necesitamos introducir la noción de *código extendido* que se presenta a continuanción [4].

Definición 3.1.4. Para cualquier código C sobre F_q, el código *extendido* $\bar{C}$ está definido por

$$\bar{C} = \{(c_1...c_n - \textstyle\sum_{i=1}^n c_i) \,:\, c_1...c_n \in C\}$$

Cuando $q = 2$ la coordenada extra $-\sum_{i=1}^n c_i = \sum_{i=1}^n c_i$ es llamada *coordenada de control de paridad.*

Teorema 3.1.5. *Si C es un código (n,M,d) sobre F_q, entonces $\bar{C}$ es un código $(n+1,M,d')$ sobre F_q con $d \leq d' \leq d+1$. Si C es lineal, entonces $\bar{C}$ también lo es. Más aún, cuando C es lineal*

$$\left(\begin{array}{ccc|c} & & & 0 \\ & H & & \vdots \\ & & & 0 \\ \hline 1 & \cdots & 1 & 1 \end{array}\right)$$

es una matriz de control de paridad de $\bar{C}$, con H una matriz de control de paridad de C.

Demostración. Es claro que la longitud de $\bar{C}$ es $n+1$ por definición de $\bar{C}$. Pasemos a demostrar que $d \leq d' \leq d+1$.
Sea $c = c_1...c_n \in C$ y $c' = c_1...c_n(-\sum_{i=1}^n c_i) \in \bar{C}$. Veamos que sucede con el peso de cada palabra código de $\bar{C}$.

(a) c' está construido desde $c \in C$ con $w(c) > d$. Luego

$$\text{si } -\sum_{i=1}^{n} c_i = 0 \text{ entonces } w(c') = w(c) > d$$
$$\text{si } -\sum_{i=1}^{n} c_i \neq 0 \text{ entonces } w(c') = w(c) + 1 > d + 1$$

Con esto concluimos que si c' está construido desde $c \in C$ con $w(c) > d$ entonces $w(c') \geq d + 1$.

(b) c' está construido desde $c \in C$ con $w(c) = d$. Luego

$$\text{si } -\sum_{i=1}^{n} c_i = 0 \text{ entonces } w(c') = w(c) = d$$
$$\text{si } -\sum_{i=1}^{n} c_i \neq 0 \text{ entonces } w(c') = w(c) + 1 = d + 1$$

Con esto concluimos que si c' está construido desde $c \in C$ con $w(c) = d$ entonces $w(c') = d$ o $w(c') = d + 1$ concluyendo así que $d \leq d' \leq d + 1$.

Para demostrar que si C es lineal entonces $\bar{C}$ también lo es, sean $a' = a_1...a_n(-\sum_{i=1}^{n} a_i)$, $b' = b_1...b_n(-\sum_{i=1}^{n} a_i) \in \bar{C}$ con $a_1...a_n,\ b_1...b_n \in C$ y $\lambda \in F_q$.

$$\begin{aligned} a' + \lambda b' &= a_1...a_n(-\sum_{i=1}^{n} a_i) + \lambda b_1...b_n(-\sum_{i=1}^{n} b_i) \\ &= a_1 + \lambda b_1 \quad ... \quad a_n + \lambda b_n \quad \left[\left(-\sum_{i=1}^{n} a_i\right) + \lambda\left(-\sum_{i=1}^{n} b_i\right)\right] \\ &= a_1 + \lambda b_1 \quad ... \quad a_n + \lambda b_n \quad \left(-\sum_{i=1}^{n} a_i + \lambda b_i\right) \end{aligned}$$

y como la upla $a_1 + \lambda b_1$ ... $a_n + \lambda b_n \in C$ por ser C lineal, concluimos que $a' + \lambda b' \in \bar{C}$ y además $\bar{C}$ contiene el elemento nulo. Así $\bar{C}$ es un código lineal.
Por último, sea $c' = c_1...c_n(-\sum_{i=1}^{n} c_i) \in \bar{C}$, con $c_1...c_n \in C$. Si hacemos

$$H' = \left(\begin{array}{ccc|c} & & & 0 \\ & H & & \vdots \\ & & & 0 \\ \hline 1 & \cdots & 1 & 1 \end{array}\right)$$

es fácil probar que $c'H'^{\top} = 0$ (ejercicio para el lector). Como esto se cumple para todo $c' \in \bar{C}$ y notando que las filas de H' son linealmente independientes, se concluye que H' es una matriz de control de paridad de $\bar{C}$. □

Ejemplo 3.1.6. Sea H una matriz de control de paridad del código $[7, 4, 3]$ que llamaremos C, con

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

Así, por Teorema (3.1.5), una matriz de control de paridad de $\bar{C}$ es

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

Se deja al lector verificar que la segunda, tercera, cuarta y quinta columna de esta matriz son linealmente dependientes y que cualquier terna de columnas son linealmente indepentientes. Luego por Corolario (2.3.16), la distancia mínima de $\bar{C}$ es 4.

Definición 3.1.7. Para cualquier código C sobre F_q, el código *perforado* en la i-ésima coordenada de C denotado por C^* está definido por

$$C^* = \{(c_1...c_{i-1}c_{i+1}c_n \ : \ c_1...c_n \in C, 1 \leq i \leq n\}$$

Entonces, sea C un código sobre F_q y C^* el código perforado en la i-ésima coordenada de C, una palabra c^* es una palabra código de C^* si y sólo si c^* se obtuvo suprimiendo la i-ésima coordenada de alguna palabra c del código C.

Teorema 3.1.8. *Sea C un $[n, k, d]$ código lineal sobre F_q, con $d \geq 2$. Entonces existe un código $[n-1, k, d-1]$ sobre F_q.*

Demostración. Como C tiene distancia mínima d, existe $c \in C$ tal que $wt(c) = d$. De la palabra código c elegimos una coordenada donde c tenga un valor no nulo. Suprimimos en todas las palabras código la coordenada elegida con anterioridad. Formamos así el código C^* de longitud $n-1$ y distancia mínima $d-1$.
Veamos que C^* es lineal. Es claro que C^* contiene al elemento nulo. Supongamos que la coordenada eliminada es la i-ésima, entonces sean
$a^* = a_1...a_{i-1}a_{i+1}...a_n$, $b^* = b_1...b_{i-1}b_{i+1}...b_n \in C^*$ con $a = a_1...a_n$, $b = b_1...b_n \in C$ y $\lambda \in F_q$. Veamos que

$$\begin{aligned} a^* + \lambda b^* &= a_1...a_{i-1}a_{i+1}...a_n + \lambda b_1...b_{i-1}b_{i+1}...b_n \\ &= a_1 + \lambda b_1 \quad ... \quad a_{i-1} + \lambda b_{i-1} \quad a_{i+1} + \lambda b_{i+1} \quad ... \quad a_n + \lambda b_n. \end{aligned}$$

Como $a_1 + \lambda b_1...a_{i-1} + \lambda b_{i-1}$ son las primeras $i-1$ coordenadas de $a + \lambda b$ y
$a_{i+1} + \lambda b_{i+1}...a_n + \lambda b_n$ son las últimas $n-i$ coordenadas de $a + \lambda b$, concluimos que $a^* + \lambda b^* \in C^*$ y por lo tanto C^*es lineal.
Finalmente la dimensión de C^* es k puesto que $d \geq 2$. □

Teorema 3.1.9. *Sea d un número impar positivo. Entonces*

(a) existe un código binario (n, M, d) si y sólo si existe un código binario $(n+1, M, d+1)$. Más aún $A_q(n+1, d+1) = A_q(n, d)$.

(b) existe un código binario $[n, k, d]$ si y sólo si el código binario $[n+1, k, d+1]$ existe y $B_q(n+1, d+1) = B_q(n, d)$

Demostración. (a) $\Rightarrow$) Supongamos que existe el código binario (n, M, d) llamado C, donde d es impar. Entonces $\bar{C}$ es un código $(n+1, M, d')$, con $d \leq d' \leq d+1$. Sea $c' = c_1...c_n(-\sum_{i=1}^{n} c_i) \in \bar{C}$ con $c_1...c_n \in C$. Observemos que si c' está construido desde $c = c_1...c_n \in C$ con $w(c) > d$, se tiene que $w(c') \geq d+1$. Por otro lado si c' está construido desde $c = c_1...c_n \in C$ con $w(c) = d$ se tiene que $\sum_{i=1}^{n} c_i = 1$ debido a que d es impar, por lo tanto $w(c') = d+1$ y $d' = d+1$.

$\Leftarrow$) Supongamos que existe D, el código binario $(n+1, M, d+1)$ donde d es impar. Sean $a', b' \in D$ tal que $d(a', b') = d+1$ i.e. difieren en $d+1 \geq 2$ coordenadas. Elegimos una coordenada de a' y b' difieran, y construimos D^* suprimiendo la coordenada elegida de todas las palabras código en D. Así D^* es un código binario (n, M, d).
De lo demostrado, resulta inmediato que $A_q(n+1, d+1) = A_q(n, d)$.

(b) La tésis sigue de los Teoremas (3.1.5) y (3.1.8).

□

3.2. Cotas inferiores

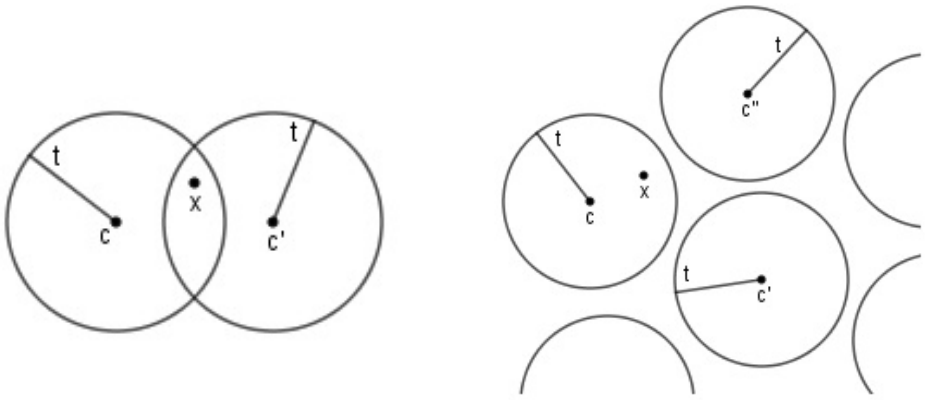
La siguiente definición resulta muy útil para continuar con la descripción del principal problema de la teoría de códigos, así como también para resignificar resultados anteriores [5].

Definición 3.2.1. Sea A un alfabeto de tamaño q, donde $q > 1$. Sea $x \in A^n$, $r \geq 0$, la *esfera de radio r y centro x* denotada como $S(x, r)$ es

$$S(x, r) = \{y \in A^n : d(x, y) \leq r\}$$

Observación 3.2.2. Con esta definición el Teorema (1.3.6), puede ser interpretado de la siguiente forma. Si $d(C) \geq 2t+1$, entonces las esferas de radio t centradas en palabras códigos de C son disjuntas [5][6]. Para demostrar esto supongamos $x \in F_q^n$ con $x \in S(c, t)$ y $x \in S(c', t)$ tal que $c, c' \in C$ (Fig.(3.1a)), entonces por desigualdad triangular se tiene que

$$d(c, c') \leq d(c, x) + d(x, c') \leq 2t$$



(a) Esferas que comparten un elemento distinto del centro.

(b) Esferas con centro una palabra código y que no comparten elementos.

Figura 3.1.

lo que contradice el hecho que $d(C) \geq 2t + 1$.

Así, si la palabra código c es enviada y se producen a lo sumo t errores recibiendo la palabra x, entonces x debe estar en $S(c, t)$ y en ninguna otra esfera, y esto nos ayuda a ver que x se decodifica como c siguiendo la decodificación por distancia mínima (Fig.(3.1b)).

Lema 3.2.3. *La cantidad de elementos que contiene una esfera de radio r en F_q^n, $0 \leq r \leq n$, es exactamente*

$$\binom{n}{0} + \binom{n}{1}(q-1) + ... + \binom{n}{r}(q-1)^r$$

Demostración. Sea x un elemento fijo de F_q^n. Calculemos cuántas palabras $y \in F_q^n$ están a una distancia exacta de 1 de x. Primero elegimos una coordenada donde y difiere de x. Esto puede hacerse de $\binom{n}{1}$ formas y esta coordenada puede tomar $q-1$ valores de F_q. Luego hay $\binom{n}{1}(q-1)$ elementos de F_q^n que están a una distancia de 1 de x.
Calculemos cuántas palabras $y \in F_q^n$ están a una distancia exacta de 2 de x. Elegimos dos coordenadas donde y difiere de x. Esto puede hacerse de $\binom{n}{2}$ formas y cada una de estas dos coordenadas puede tomar $q-1$ valores de F_q. Luego hay $\binom{n}{2}(q-1)^2$ elementos de F_q^n que están a una distancia de 2 de x.
De forma general, calculemos cuántas palabras $y \in F_q^n$ están a una distancia exacta

de m de x. Las m coordenadas en las que y difiere de x puede elegirse de $\begin{pmatrix} n \\ m \end{pmatrix}$ formas posibles y cada una de las m coordenadas puede elegirse de $q-1$ elementos de F_q. Así hay $\begin{pmatrix} n \\ m \end{pmatrix}(q-1)^m$ elementos posibles que distan exactamente en m de x.
Teniendo en cuenta que la única palabra a una distancia de 0 de x es x, se tiene que la cantidad de elementos que contiene una esfera de radio r en F_q^n, $0 \leq r \leq n$, es exactamente

$$\begin{pmatrix} n \\ 0 \end{pmatrix} + \begin{pmatrix} n \\ 1 \end{pmatrix}(q-1) + ... + \begin{pmatrix} n \\ r \end{pmatrix}(q-1)^r$$

□

Observación 3.2.4. Es claro que si $r > n$ ocurre que $S(x,r) = F_q^n$ para cualquier $x \in F_q^n$.

Teorema 3.2.5 (Sphere-covering bound)**.** *Para un entero $q > 1$ y enteros n, d tal que $1 \leq d \leq n$ se tiene que*

$$\frac{q^n}{\sum_{i=1}^{d-1} \begin{pmatrix} n \\ i \end{pmatrix}(q-1)^i} \leq A_q(n,d). \tag{3.2.1}$$

Demostración. Sea $C = \{c_1, ..., c_M\}$ un código óptimo (n, M, d) sobre el alfabeto A con $|A| = q$, i.e. $M = A_q(n,d)$.
Ya que C tiene el máximo tamaño, no puede haber una palabra en A^n que no esté en C tal que diste al menos en d de cada palabra del código C. Si así lo hiciera podríamos formar un código $(n, M+1, d)$ lo cuál es absurdo pues $A_q(n,d) = M$.
Más aún, sea $x \in A^n$, existe al menos una palabra código $c_j \in C$ tal que $d(x, c_j)$ sea a lo sumo $d-1$ i.e. $x \in S(c_j, d-1)$. Así, cada palabra de A^n está incluido en alguna esfera $S(c_j, d-1)$, con $1 \leq j \leq M$. En símbolos

$$A^n \subseteq \cup_{j=1}^{M} S(c_j, d-1).$$

Como $|A^n| = q^n$ y $|s(c_j, d-1)| = \sum_{i=1}^{d-1} \begin{pmatrix} n \\ i \end{pmatrix}(q-1)^i$
Así

$$q^n \leq M \sum_{i=1}^{d-1} \begin{pmatrix} n \\ i \end{pmatrix}(q-1)^i$$

y como $M = A_q(n,d)$

$$\frac{q^n}{\sum_{i=1}^{d-1} \begin{pmatrix} n \\ i \end{pmatrix}(q-1)^i} \leq A_q(n,d).$$

□

Teorema 3.2.6 (Cota de Gilbert-Varshmov)**.** *Sea q la potencia de un número primo. Entonces existe un código $[n, k]$ sobre F_q con distancia al menos d, suponiendo que se cumpla que*

$$\sum_{i=0}^{d-2} \left(\begin{array}{c} n-1 \\ i \end{array} \right) (q-1)^i < q^{n-k}$$

con $2 \leq d \leq n$ y $1 \leq k \leq n$.

Para demostrar este teorema construiremos una matriz H de forma tal que podamos usar el Teorema (2.3.15) para concluir la tésis.

Demostración. Denotaremos la j-ésima columna de H como c_j.
Sea $c_1 \in F_q^{n-k}$ un elemento no nulo. Sea $c_2 \in F_q^{n-k}$ tal que $c_2 \notin \langle c_1 \rangle$. Para j con $2 < j \leq n$, sea $c_j \in F_q^{n-k}$ tal que c_j no pertenece a ningún conjunto generado por un conjunto de $d-2$ elementos, o menos, del conjunto $c_1, c_2, ..., c_{j-1}$. Ahora, la cantidad de elementos en el total de los conjuntos generados por $d-2$ elementos, o menos, del conjunto $c_1, c_2, ..., c_{j-1}$ es

$$\sum_{i=0}^{d-2} \left(\begin{array}{c} j-1 \\ i \end{array} \right) (q-1)^i$$

y por hipótesis se cumple

$$\sum_{i=0}^{d-2} \left(\begin{array}{c} j-1 \\ i \end{array} \right) (q-1)^i \leq \sum_{i=0}^{d-2} \left(\begin{array}{c} n-1 \\ i \end{array} \right) (q-1)^i < q^{n-k}$$

por lo tanto siempre es posible encontrar encontrar c_j con $2 \leq j \leq n$ y por lo tanto construir H tal que cualquier conjunto de $d-1$ columnas son linealmente independientes. El complemento ortogonal de H es un código $[n, k]$ y por Teorema (2.3.15) tiene distancia al menos d. □

Corolario 3.2.7. *Sea q la potencia de un número primo. Entonces $A_q(n, d) \geq q^{k_1}$, donde k_1 es el valor más grande de k tal que*

$$q^k < \frac{q^n}{\sum_{i=0}^{d-2} \left(\begin{array}{c} n-1 \\ i \end{array} \right) (q-1)^i}$$

Demostración. Que k_1 es el valor más grande de k tal que

$$q^k < \frac{q^n}{\sum_{i=0}^{d-2} \left(\begin{array}{c} n-1 \\ i \end{array} \right) (q-1)^i}$$

es equivalente a decir que k_1 es el valor más grande de k tal que

$$\sum_{i=0}^{d-2} \left(\begin{array}{c} n-1 \\ i \end{array} \right) (q-1)^i < q^{n-k}$$

Así existe el código $[n, k_1]$ con distancia al menos d. Este código tiene q^{k_1} elementos lo que significa que

$$q^{k_1} \leq B_q(n, d) \leq A_q(n, d).$$

□

3.3. Cotas superiores y códigos perfectos

A continuación se presenta una cota superior que permite clasificar ciertos códigos llamados perfectos. En esta sección se describe dos códigos perfectos: el código Hamming, y el código Golay.

Teorema 3.3.1 (Cota de Hamming)**.** *Para un entero $q > 1$ y enteros n y d tales que $1 \leq d \leq n$ se cumple*

$$A_q(n, d) \leq \frac{q^n}{\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i} (q-1)^i}$$

Demostración. Sea $C = \{c_1, c_2, ..., c_M\}$ un código óptimo (n, M, d) sobre A, un alfabeto de tamaño q, i.e. $M = A_q(n, d)$. Sea $t = \frac{d-1}{2}$, entonces las esferas $S(c_i, t)$ con $i = 1, ..., M$ son disjuntas y se tiene que

$$\cup_{i=1}^{M} S(c_i, t) \subseteq A^n$$

Esta unión es una unión disjunta.
Como $|A^n| = q^n$ y $|S(c_i, t)| = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$ entonces se cumple que

$$M \cdot \left(\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \right) \leq q^n$$

$$M = A_q(n, d) \leq \frac{q^n}{\sum_{i=0}^{t} \binom{n}{i} (q-1)^i}$$

□

Definición 3.3.2. Un código sobre un alfabeto de q elementos tal que sus parámetros haga que se dé la igualdad en la cota de Hamming i.e. si tiene exactamente $q^n / \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i} (q-1)^i$ palabras código, es llamado *código perfecto.*

Un código perfecto (n, M, d) sobre el alfabeto A, con $|A| = q$, por definición satisface que

$$M \cdot \left(\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i} (q-1)^i \right) = q^n$$

esto quiere decir que $\cup_{j=1}^{M} S(c_j, (d-1)/2) = A^n$, donde c_j con $j = 1...M$ son las M palabras código y la unión es una unión de conjuntos disjuntos al ser el radio de cada esfera $(d-1)/2$. (Fig.(3.2))

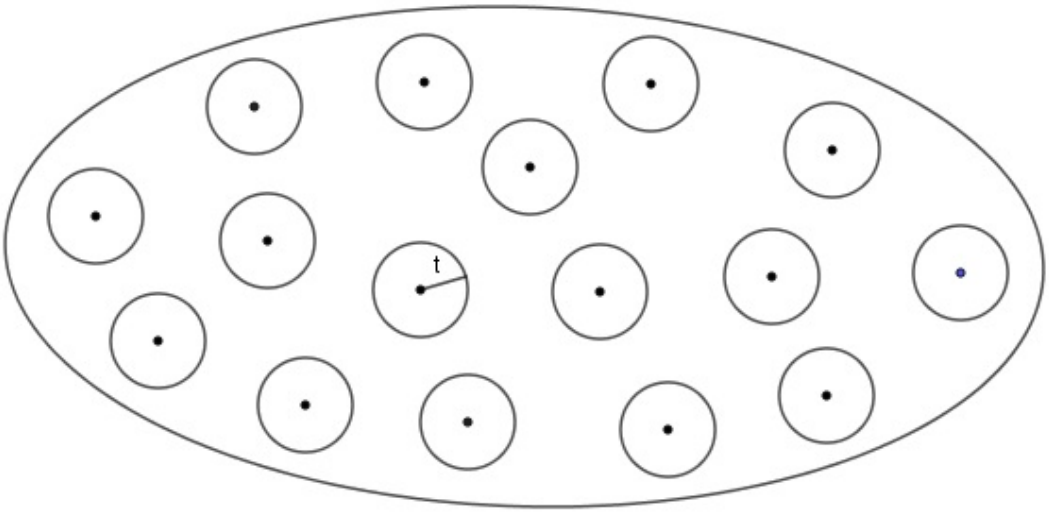


Figura 3.2.: En un código perfecto no existen elementos fuera de una esfera de centro una palabra código y radio $t = (d-1)/2$.

Ejemplo 3.3.3. Los siguientes códigos son llamados códigos perfectos triviales.

(a) El código lineal $C = F_q^n$ $(d = 1)$.

(b) Cualquier código C, tal que $|C| = 1$.

(c) Los códigos de repetición de F_2^n con n impar, que consisten en únicamente dos palabras y tienen distancia mínima n.

Los códigos Hamming y Golay que brevemente describiremos a continuación, son códigos perfectos no triviales.

3.3.1. Código Hamming

3.3.1.1. Código Hamming binario

Definición 3.3.4. Sea $r \geq 2$. Un código lineal binario de longitud $n = 2^r - 1$, con H una matriz de control de paridad cuyas columnas consisten en todos los vectores no nulos de F_2^r, es llamado *código (binario) Hamming* y se lo denota como $Ham(r, 2)$.

Observación 3.3.5. 1. Las r filas de H son linealmente independientes ya que todos los vectores de peso uno de F_2^r son columnas de H.

2. Como la definición no establece un orden para las columnas de H, para un r fijo, $Ham(r,2)$ es único salvo equivalencias, esto quiere decir que todos los códigos $Ham(r,2)$ son equivalentes [4][5].

Ejemplo 3.3.6. (a) Para $r = 2$ una matriz de control de paridad es

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

Esta matriz es una matriz de control de paridad del código de repetición $[3, 1]$

(b) Para $r = 3$ una matriz de control de paridad es

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

Si permutamos las columnas de H formando la matriz H_1, una matriz de control de paridad en forma estándar, podemos determinar G, una matriz generatriz estandar del código.

$$H_1 = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

El código $Ham(3,2)$ es entonces el código perfecto $[7, 4, 3]$.

(c) Para $r = 4$ una matriz de control de paridad es

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

El teorema que sigue a continuación nos dirá que el código $Ham(4,2)$ es el código perfecto $[15, 11, 3]$.

Observación 3.3.7. Para un r fijo, los elementos no nulos de F_2^r representan los primeros $r-1$ naturales en el sistema binario, así una forma de ordenar las columnas de la matriz de control de paridad de $Ham(r,2)$ es hacerlo siguiendo el orden de los números naturales en forma creciente.

Teorema 3.3.8. *Sea $r \geq 2$. El código $Ham(r,2)$*

(a) tiene dimensión $k = 2^r - 1 - r$

(b) tiene distancia $d = 3$

(c) es un código perfecto.

Demostración. (a) Como H tiene r filas y $2^r - 1$ columnas, $Ham(r,2)^{\perp}$ es un código $[2^r - 1, r]$, entonces por Teorema (2.3.6) $Ham(r,2)$ es un código $[2^r - 1, 2^r - 1 - r]$.

(b) Sea H una matriz de control de paridad del código $Ham(r,2)$. Como las columnas de H son todas los elementos no nulas de F_2^r, cualquier par de columnas de H son linealmente independientes. Al ser $r \geq 2$, la representación binaria de uno, dos y tres forman parte de las columnas de H y éstas son linealmente dependientes. Por Corolario (2.3.16) la distancia de $Ham(r,2) = 3$, y por lo tanto, $Ham(r,2)$ corrige un único error.

(c) Para determinar si $Ham(r,2)$ es un código perfecto deberemos ver que sucede con la cota de Hamming. Como $M = 2^{2^r-1-r}$ se sigue que

$$2^{2^r-1-r} \sum_{i=0}^{1} \begin{pmatrix} 2^r - 1 \\ i \end{pmatrix} (2-1)^i = 2^{2^r-1-r}(1 + 2^r - 1) = 2^{2^r-1}$$

Por lo tanto $Ham(r,2)$ es un código perfecto.

□

3.3.1.2. Decodificación del codigo Hamming binario

Dijimos que $Ham(r,2)$ tiene distancia 3. Esto quiere decir que las palabras de peso uno de $F_2^{2^r-1}$ no son elementos de $Ham(r,2)$. ¿Pueden haber dos palabras de peso uno en una misma coclase?
Sea $C_r = Ham(r,2)$. Recordemos que dos elementos $x, y \in F_2^{2^r-1}$ están el la misma coclase si y solo si $x - y \in C_r$. Así la respuesta a la pregunta de arriba es no, no puede haber dos palabras de peso uno en una misma coclase ya que su diferencia es una palabra de peso dos y no pertenece a C_r. Luego estos elementos serán los representantes de cada coclase.
Sea $e_i \in F_2^{2^r-1}$, con $i = 1, ..., 2^r - 1$ la palabra de peso uno, que tiene su i-ésima coordenada no nula y H una matriz de control de paridad de $Ham(r,2)$. El lector puede verificar que $S(e_i)$ es la i-ésima columna de H.
Suponiendo que las columnas de H fueron ordenadas de manera creciente (de acuerdo a la representación binaria de los primeros $r-1$ números naturales), la decodificación de $Ham(r,2)$ está dado por los siguientes pasos:

Paso 1 Si se recibió la palabra y, calculamos $S(y)$.

Paso 2 Si $S(y) = 0$, entonces asumimos que y es la palabra código enviada.

Paso 3 Si $S(y) \neq 0$, asumimos que ocurrió un único error. $S(y)$ será la representación binaria de la coordenada donde se produjo el error y solo resta corregirla.

Observación 3.3.9. Cuando $S(y) \neq 0$ se asume que ha ocurrido un error, ya que si ocurren dos o más errores se decodificará como la palabra código incorrecta.

Ejemplo 3.3.10. Consideremos $Ham(3,2)$, el código $[7,4]$ del ejemplo anterior y su matriz de control de paridad H. Supongamos que la palabra código $c = 0111100$ fue enviada y se recibió: a) $y = 0111000$ b) $y = 0111001$. En cada caso decodificar y.

a) $0111000 \cdot H^\top = 101$. 101 es el número 5 en binario. Luego el error se cometió en la quinta coordenada, por lo que y se decodifica como $0111100 = c$.

b) $0111001 \cdot H^\top = 010$. 010 es el numero 2 en binario. Luego el error se cometió en la segunda coordenada, por lo que y se decodifica como $0011001 \neq c$.

3.3.1.3. Código Hamming q-ario

Sea $q \geq 2$ la potencia de un número primo. Nos interesa saber cuántos subespacios de dimensión uno existen en F_q^r con $r \geq 2$. Sea $x \in F_q^r$ un vector un elemento no nulo, se tiene que $\langle x \rangle$ tiene exactamente $q-1$ elementos no nulos de F_q^r. Cada subespacio de dimensión uno tiene entonces $q-1$ elementos no nulos, y así la cantidad de subespacios de dimensión uno en F_q^r es de $\dfrac{q^r-1}{q-1}$.

Definición 3.3.11. Sea $r \geq 2$. Un código lineal q-ario, con una matriz de control de paridad H cuyas columnas consisten en vectores no nulos, uno de cada subespacio de dimensión uno de F_q^r, es llamado *código Hamming* y se denota $Ham(r,q)$.

Observación 3.3.12.

1. Como la cantidad de subespacios de dimensión uno en F_q^r es de $\dfrac{q^r-1}{q-1}$, la longitud del código es $\dfrac{q^r-1}{q-1}$.
2. Notemos que los subespacios de dimensión uno de F_2^r son los conjuntos unitarios formados por cada elemento no nulo de F_2^r, de ahí la definición de código Hamming binario.

Ejemplo 3.3.13. (a) Para $r = 2$ y $q = 3$ una matriz de control de paridad H para $Ham(2,3)$ es

$$H = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

(b) Para $r = 2$ y $q = 5$ una matriz de control de paridad H_1 para $Ham(2,5)$ es

$$H_1 = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}$$

(c) Para $r = 3$ y $q = 3$ una matriz de control de paridad H_2 para $Ham(3,3)$ es

$$H_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{pmatrix}$$

Teorema 3.3.14. *Sea $r \geq 2$ y $q \geq 2$ con q una potencia prima. El código $Ham(r,q)$*

(a) es un código $\left[\dfrac{q^r-1}{q-1}, \dfrac{q^r-1}{q-1} - r, 3\right]$,

(b) es un código perfecto.

Demostración. (a) Sea H una matriz de control de paridad de $Ham(r,q)$, i.e. H es una matriz generadora del código $Ham(r,q)^{\perp}$. H tiene tamaño $r \times \dfrac{q^r-1}{q-1}$ y así $Ham(r,q)^{\perp}$ es un código $\left[\dfrac{q^r-1}{q-1}, r\right]$, entonces por Teorema (2.3.6) $Ham(r,q)$ es un código $\left[\dfrac{q^r-1}{q-1}, \dfrac{q^r-1}{q-1} - r\right]$.

Como cada columna es un elemento de cada uno de los $\dfrac{q^r-1}{q-1}$ subespacios de dimensión uno de F_q^r, cualquier par de columnas de H son linealmente independientes. Supongamos, sin pérdida de generalidad, que una de las cólunas de H tiene a 1 en la primera cordenada y cero en las restantes, otra con 1 en la segunda coordenada y cero en las restantes y una tercera con 1 en la primera y segunda coordenada y cero en las restantes. Está claro que estas columnas cumplen con estar en subespacios de dimensión uno distintos y se observa que son linealmente dependientes. Así por Corolario (2.3.16) se tiene que $d(Ham(r,q)) = 3$.

(b) Veamos que

$$\begin{aligned} q^{\frac{q^r-1}{q-1}-r} \sum_{i=0}^{1} \binom{\frac{q^r-1}{q-1}}{i} (q-1)^i &= q^{\frac{q^r-1}{q-1}-r}(1 + \frac{q^r-1}{q-1}(q-1)) \\ &= q^{\frac{q^r-1}{q-1}-r}(q^r) \\ &= q^{\frac{q^r-1}{q-1}} \end{aligned}$$

y así $Ham(r,q)$ es perfecto.

□

3.3.1.4. Decodificación del código Hamming q-ario

Ya que $Ham(r,q)$ es un código perfecto 1-error-corrector, los representantes de las coclases serán justamente, además de la palabra código 0, elementos de peso 1 de

F_q^n, con $n = \dfrac{q^r - 1}{q - 1}$. Sea $e_{j,b}$, con $1 \leq j \leq n$ y $b \in F_q - \{0\}$, el vector cuya j-ésima coordenada es b y el resto cero. El síndrome de $e_{j,b}$ es

$$S(e_{j,b}) = e_{j,b}H^\top = b \cdot c_j^\top$$

donde H es una matriz de control de paridad de $Ham(r, q)$ y c_j es la j-ésima columna de H. El esquema de decodificación es el siguiente.

Paso 1 Si se recibió la palabra y, calculamos $S(y)$.

Paso 2 Si $S(y) = 0$ asuminos que no se produjeron errores.

Paso 3 Si $S(y) \neq 0$, entonces encontrar $e_{j,b}$ tal que $S(y) = S(e_{j,b})$. Se decodifica y como $y - e_{j,b}$

Ejemplo 3.3.15. (a) Consideremos $Ham(2, 5)$ y la matriz de control de paridad H_1 del Ejemplo (3.3.13). Supongamos que se recibió la palabra $y = 203031$. Como $S(y) = 2\ 3 = 2(1\ 4)$. Así y se decodifica como $y - e_{6,2} = 203031 - 000002 = 203034$.

(b) Consideremos $Ham(3, 3)$ y la matriz de control de paridad H_2 del Ejemplo (3.3.13). Supongamos que se recibió la palabra a) $y = 2100111220012$. b) $y = 1001012222221$.

a) $S(y) = 120 = S(e_{11,1})$ (i.e. es la onceava columna de H_2). Así y se decodifica como $y - e_{11,1} = 2100111220012 - 0000000000100 = 2100111220212$.

b) $S(y) = 220 = 2(110) = S(e_{8,2})$. Así y se decodifica como $y - e_{8,2} = 1001012222221 - 0000000200000 = 1001012122221$.

En resumen, para $Ham(r, 2)$ y $Ham(r, q)$ se tiene que

$$Ham(r, 2),\ r \geq 2$$

Longitud	$2^r - 1$
Dimensión	$2^r - 1 - r$
$d(Ham(r, 2))$	3

Es un código perfecto que corrigue sólo error, único salvo equivalencias. Las columnas de H son todos los elementos no nulos de F_2^r.

$Ham(r,q),\ r \geq 2,\ q \geq 2$	
Longitud	$\dfrac{q^r-1}{q-1}$
Dimensión	$\dfrac{q^r-1}{q-1} - r$
$d(Ham(r,2))$	3

Es un código perfecto que corrige sólo error, único, salvo equivalencias.

3.3.2. Código Golay

Aquí se presenta el código Golay por medio de una matriz generadora, tal como lo hizo el mismo Golay en una de sus publicaciones en 1949. Éste planteó el problema de encontrar todos los códigos perfectos, sin embargo no pudo ser resuelto por él, pero sí por J. H. van Lint y A. Tietäväinen en 1973 (aunque solo para casos que sean sobre alfabetos de tamaño la potencia de un número primo) [4].
Para continuar necesitaremos definir una nueva operación entre dos palabras en F_2^n [3].

Definición 3.3.16. Sea $x, y \in F_2^n$. Si $x = x_1x_2...x_n$ y $y = y_1y_2...y_n$, entonces $x \cap y$ está definido por

$$x \cap y = x_1 \cdot y_1\ x_2 \cdot y_2...x_n \cdot y_n,\ \text{así}\ x \cap y \in F_2^n.$$

Lema 3.3.17. *Sea $x, y \in F_2^n$, entonces*

(a) $d(x,y) = wt(x) + wt(y) - 2wt(x \cap y)$.

(b) $wt(x \cap y) \equiv x \cdot y\ (mod\ 2)$.

Demostración. (a) Sabemos que con $x, y \in F_2^n$, por Lema (2.1.4), se cumple que $d(x,y) = wt(x-y) = wt(x+y)$, esta última igualdad se debe a que estamos trabajando sobre F_2.
Es claro que $wt(x+y) = wt(x)+wt(y)$ es falso. Sabemos que $wt(x+y)$ cuenta la cantidad de coordenadas en donde x e y difieren que son las coordenadas donde $x+y$ tiene $1s$; también que $wt(x)+wt(y)$ cuenta la cantidad de $1s$ total en x e y, i.e. las coordenadas en donde x e y difieren y donde las que no difieren y contienen $1s$. Éstas últimas se cuentan dos veces. Por lo tanto para hacer verdadera la igualdad debemos "quitarlas"de la cuenta. Ahora, $x \cap y$ tiene $1s$ en las coordenadas donde x e y no difieren. Por lo tanto $2wt(x \cap y)$ es lo que se debe "quitar". De esto y del hecho que $d(x,y) = wt(x+y)$ se sigue que

$$d(x,y) = wt(x) + wt(y) - 2wt(x \cap y)$$

(b) Sabemos que

$$x \cap y = x_1 \cdot y_1 \; x_2 \cdot y_2 ... x_n \cdot y_n \text{ y } x \cdot y = \sum_{i=1}^{n} x_i \cdot y_i$$

Tenemos que:

- $wt(x \cap y)$ es par, $x \cdot y = \sum_{i=1}^{n} x_i \cdot y_i = 0$ y así

$$wt(x \cap y) \equiv x \cdot y (\text{mod} 2)$$

- $wt(x \cap y)$ es impar, $x \cdot y = \sum_{i=1}^{n} x_i \cdot y_i = 1$ y así

$$wt(x \cap y) \equiv x \cdot y (\text{mod} 2)$$

□

Recordemos que el primer inciso del Teorema (3.1.9) dice que existe un código binario (n, M, d) si y sólo si existe un código binario $(n+1, M, d+1)$ con d impar. Definimos primero el código Golay binario extendido para luego a través de el teorema mencionado anteriormente mostrar las propiedades del código Golay binario.

Definición 3.3.18. Sea G la siguiente matriz de tamaño 12×24

$$G = \left(\begin{array}{c|cccccccccccc} & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ I_{12} & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

donde I_{12} es la matriz identidad de orden 12. El código binario con matriz generadora G es llamado *código Golay binario extendido* y se denota como G_{24}.

Nota 3.3.19. La matriz G, la escribiremos como $G = (I_{12}|A)$ donde A tiene las últimas 12 columnas de G en ese orden.
El código G_{24} fue usado en las misiones Voyager I y Voyager II durante 1979-1981 para la transmisión a la Tierra de imágenes a color de Júpiter y Saturno, proporcionando corrección de errores en la transmisión [8].

Las demostraciones de los primeros dos lemas, de los cuatro que se enuncian a continuación, queda como ejercicio para el lector, quien puede además encontrar una forma de demostrarlos en [5].

Lema 3.3.20. *$G_{24}^{\perp} = G_{24}$ i.e. G_{24} es auto dual [5].*

Lema 3.3.21. *$(A|I_{12})$ es también una matriz generadora de G_{24}.*

Lema 3.3.22. *Cada palabra código de G_{24} tiene peso divisible por 4.*

Demostración. Si $x, y \in G_{24}$ entonces $x \cdot y = 0$ por ser G_{24} auto dual y como $wt(x \cap y) \equiv x \cdot y (\text{mod} 2)$, $wt(x \cap y)$ es par.
Sean x e y dos filas de G. Por Lema (2.1.4) y Lema (3.3.17)

$$wt(x + y) = wt(x) + wt(y) - 2wt(x \cap y)$$

Observemos que el peso de cada fila de G es divisible por 4. Así para cualquier par de filas de G, los dos primeros términos de la igualdad anterior son divisibles por 4. Dijimos que $wt(x \cap y)$ es par, por lo tanto el tercer término de la igualdad dada también es divisible por 4. Luego $wt(x + y)$ es divisible por 4.
Sea z una fila de G y w la combinación lineal de dos filas de G. Como vimos anteriormente $wt(w)$ es divisible por 4, dicho esto y sabiendo que $wt(z)$ también es divisible por 4, se tiene que $wt(z + w)$ es divisible por 4.
Siguiendo este razonamiento, cualquier combinacion lineal de las filas de G es divisible por 4, i.e. cualquier palabra código es divisible por 4. □

Lema 3.3.23. *G_{24} no tiene palabras código de peso 4.*

Demostración. Escribamos una palabra código $x = x_1x_2...x_{24}$ como $(L|R)$ donde $L = x_1x_2...x_{12}$ y $R = x_{13}x_{14}...x_{24}$.
Supongamos que x es una palabra código de peso 4. Entonces uno de los siguientes casos debería ocurrir:

Caso 1 $wt(L) = 0$, $wt(R) = 4$. Esto no puede ocurrir pues G está en forma estándar y $wt(L) = 0$ ocurre únicamente cuando $x = 0$.

Caso 2 $wt(L) = 1$, $wt(R) = 3$. Que $wt(L) = 1$ significa que x es una de las filas de G de lo cual se sigue que $wt(R) \neq 3$, así este caso no puede ocurrir.

Caso 3 $wt(L) = 2$, $wt(R) = 2$. x es la suma de 2 filas de G. Queda para el lector verificar que ninguna suma entre dos filas de G hace que $wt(R) = 2$.

Caso 4 $wt(L) = 3$, $wt(R) = 1$. x es suma de 3 filas de G. Por Lema (3.3.21) $(A|I_{12})$ genera exactamente el mismo código G_{24}, luego x debería ser una de las filas de $(A|I_{12})$ donde resulta que $wt(L) \neq 3$.

Caso 5 $wt(L) = 4$, $wt(R) = 0$. Usando nuevamente $(A|I_{12})$ como matriz generatriz de G_{24}, $wt(R) = 0$ ocurre siempre que $x = 0$.

Así como no ocurre ninguno de estos casos, concluimos que no existe ninguna palabra código de peso 4. □

Teorema 3.3.24. *G_{24} es un código* $[24, 12, 8]$.

Demostración. Lo que hay que probar es que $d = 8$. Observemos que tiene al menos una fila de peso 8. Luego por el Lema (3.3.22) y el Lema (3.3.23) se sigue que $d = 8$. □

Definición 3.3.25. Sea $\hat{G}$ la siguiente matriz de tamaño 12×23

$$\hat{G} = \left(I_{12}|\hat{A}\right),$$

donde I_{12} es la matriz identidad de tamaño 12×12 y $\hat{A}$ es la matriz de tamaño 12×11 obtenida de la matriz A suprimiendo su última columna. El código lineal binario que tiene a $\hat{G}$ como una matriz generadora, es llamado *código Golay binario* y se denota como G_{23}.

Teorema 3.3.26. *(a) La extensión del código G_{23} es G_{24}.*

(b) La distancia de G_{23} es $d = 7$.

(c) El código G_{23} es un código perfecto 3-error-corrector.

Demostración. (a) Sea $\hat{G}$ la matriz generadora de G_{23} dada arriba y $r_i = f_{i1}...f_{in}$ con $1 \leq i \leq 12$ sus filas.
Sea $\bar{G}$ la siguiente matriz

$$\bar{G} = \begin{pmatrix} \bar{r_1} \\ \vdots \\ \bar{r_{12}} \end{pmatrix}$$

donde $\bar{r_i} = f_{i1}...f_{in}(-\sum_{j=1}^{n} f_{ij})$ con $1 \leq i \leq 12$.
Sea H una matriz de control de paridad de G_{23}. Sabemos que $\bar{H}$ dada por

$$\bar{H} = \left(\begin{array}{ccc|c} & & & 0 \\ & H & & \vdots \\ & & & 0 \\ \hline 1 & \cdots & 1 & 1 \end{array}\right)$$

es la matriz de control de paridad de la extensión de G_{23}.
Es claro que

$$\begin{pmatrix} \bar{r_1} \\ \vdots \\ \bar{r_{12}} \end{pmatrix} \cdot \left(\begin{array}{ccc|c} & & & 1 \\ & H^{\top} & & \vdots \\ & & & 1 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) = 0$$

Esto quiere decir que $\bar{G}$ es una matriz generatriz de la extensión de G_{23}. Notemos que $\bar{G}$ es la matriz generatriz de G_{24} en la Definición (3.3.18). De esto sigue la tésis.

(b) Sea c una palabra código no nula de G_{24}, entonces $wt(c) \geq 8$. Como las palabras código de G_{23} se obtienen suprimiendo la última coordenada de las palabras código de G_{24}, sea $c' \in G_{23}$ la palabra código que se obtuvo desde $c \in G_{24}$. Se tiene que $wt(c') \geq 7$ y observemos que suprimiendo la última coordenada de la cuarta fila de G_{24} resulta una palabra código de peso 7. Sigue la tésis [9].

(c) G_{23} es un código perfecto ya que

$$\begin{aligned} 2^{12} \cdot \left(\sum_{i=0}^{3} \binom{n}{i} (2-1)^i \right) &= 2^{12} \cdot \left(\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} \right) \\ &= 2^{12} \cdot (1 + 23 + 253 + 1771) \\ &= 2^{12} \cdot 2048 \\ &= 2^{12} \cdot 2^{11} \\ &= 2^{23}. \end{aligned}$$

Es 3-error-corrector ya que su distancia mínima es 7.

□

Definición 3.3.27. EL *código Golay ternario extendido*, denotado por G_{12}, es el código lineal ternario con matriz generadora

$$G = \left(\begin{array}{c|cccccc} & 0 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 1 & 2 & 2 & 1 \\ I_6 & 1 & 1 & 0 & 1 & 2 & 2 \\ & 1 & 2 & 1 & 0 & 1 & 2 \\ & 1 & 2 & 2 & 1 & 0 & 1 \\ & 1 & 1 & 2 & 2 & 1 & 0 \end{array} \right)$$

Definición 3.3.28. El *código Golay ternario* G_{11} es el código que se obtiene al suprimir la última coordenada de las palabras código de G_{12}.

Teorema 3.3.29. *G_{11} es un código $[11, 6, 5]$ y además es perfecto.*

3.3.2.1. Decodificación de G_{24}

Previo a presentar el algoritmo de decodificación, realizaremos observaciones de supuestos a considerar para realizar la decodificación; los mismos fundamentan los pasos del algoritmo de decodificación [8][10].
En la Definición (3.3.18) presentamos a $G = (I_{12}|A)$ como matriz generadora de G_{24}. El lector ya habrá notado que la matriz A es simétrica. Aún más, esta matriz cumple con que $A^{-1} = A^{\top} = A$.
Sabemos además que G_{24} es auto dual, así $G = (I_{12}|A)$ es también matriz de control de paridad de G_{24} y es la que usaremos, bajo el nombre de H, para calcular el

síndrome de las palabras recibidas.
Denotamos con A_i, para $1 \leq i \leq 12$, la i-ésima fila de A y con $a^{(i)}$ el elemento de F_2^{12} cuya i-esima componente es la única no nula. Es claro que

$$A_i = a^{(i)} \cdot A$$

Sea $e = (r, t)$ el error que se produjo en la transmisión del mensaje, en donde $r, t \in F_2^{12}$. Supongamos que se transmite una palabra x, ocurren errores que se representa con el vector error e y se ha recibido la palabra y. Recordemos que el síndrome s de y es igual al síndrome de e. Veamos que

$$\begin{aligned} s &= e \cdot H^\top \\ &= (r, t) \cdot \begin{pmatrix} I_{12} \\ A \end{pmatrix} \\ &= r \cdot I_{12} + t \cdot A \end{aligned}$$

Teniendo en cuenta que $A = A^{-1}$ se tiene que

$$t = (s + r)A \tag{3.3.1}$$

G_{24} es un código que corrige a lo sumo 3 errores, por lo tanto analizamos los casos en donde $wt(e) \leq 3$. De este modo tenemos los siguientes casos:

Caso 1 $wt(t) = 0$ y $wt(r) \leq 3$
Así $s = r$ y $e = (r, 0)$ con $wt(s) \leq 3$.

Caso 2 $wt(t) = 1$ y $wt(r) \leq 2$
Se tiene que $s = r + a^{(i)} \cdot A = r + A_i$ y así $e = (s + A_i, a^{(i)})$ con $wt(s + A_i) \leq 2$.

Caso 3 $wt(t) = 2$ o $wt(t) = 3$ y $wt(r) = 0$
Como $wt(r) = 0$, $r = 0$. Luego $s = t \cdot A$, por lo que $t = s \cdot A$ y así $e = (0, s \cdot A)$ con $2 \leq wt(t \cdot A) \leq 3$.

Caso 4 $wt(t) = 2$ y $wt(r) = 1$
Bajo este supuesto se tiene que

$$\begin{aligned} s &= a^{(i)} + t \cdot A \\ t &= (s + a^{(i)})A \\ t &= s \cdot A + A_i \end{aligned}$$

por lo tanto $e = (a^{(i)}, s \cdot A + A_i)$, con $wt(s \cdot A + A_i) = 2$

Teniendo en cuenta la notación dada anteriormente, presentamos el algoritmo de decodificación. Supongamos que se recibió la palabra y

Paso 1 Calcular $S(y) = s$.

Paso 2 Si $wt(s) \leq 3$ entonces $e = (s, 0)$. Ir al Paso 8.

Paso 3 Si $wt(s + A_i) \leq 2$ para algún $1 \leq i \leq 12$, entonces $e = (s + A_i, a^{(i)})$. Ir al Paso 8.

Paso 4 Calcular $s \cdot A$.

Paso 5 Si $wt(s \cdot A) \leq 3$, entonces $e = (0, s \cdot A)$. Ir al Paso 8.

Paso 6 Si $wt(s \cdot A + A_i) \leq 2$ para algún $1 \leq i \leq 12$ entonces $e = (a^{(i)}, s \cdot A + A_i)$. Ir al Paso 8.

Paso 7 Si no se puede construir e siguiendo los pasos anteriores, pedir una retransmisión.

Paso 8 decodificar y como $x = y + e$.

Ejemplo 3.3.30. Se recibe la palabra

$$y = \ 1 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 1 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 1$$

- Calculamos $S(y)$. $s = S(y) = \ 1 \ \ 1 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 1 \ \ 1 \ \ 1 \ \ 0 \ \ 1$
- Para $i = 6$, se tiene que $s + A_6 = \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 0$.
  Así $e = \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0$.
- Finalmente y se decodifica como $c = y + e$ i.e.

  $$e = \ 1 \ \ 1 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 1 \ \ 0 \ \ 0 \ \ 1 \ \ 0 \ \ 1 \ .$$

3.3.2.2. Decodificación de G_{23}

La decodificación de G_{23} se ralaciona a cómo lo definimos a partir G_{24}. Sea $z \in F_2^{23}$, definimos $z0$ y $z1$ como los elementos de F_2^{24} que resultan de haber agregado 0 en la última coordenada de z cuando $wt(z)$ es par, o agregar 1 en la última coordenada de z cuando $wt(z)$ es impar, i.e. lo que agregamos es la coordenada de control de paridad.
Supongamos que se recibió la palabra y. Los siguientes pasos corresponden al algorítmo de decodificación de G_{23}:

Paso 1 Calcular $wt(y)$ y formar $y0$ o $y1$ segun corresponda.

Paso 2 Decodificar yi (con $i = 0$ o $i = 1$) como c de acuerdo a la decodificación dada para G_{24}.

Paso 3 Suprimir la última coordenada de c.

Resumidamente podemos decir que

El código Golay

	G_{23}	G_{24}	G_{11}
Longitud	23	24	11
Dimensión	12	12	6
Distancia mínima	7	8	5
Errores que corrige	3	3	2

G_{23} y G_{11} son códigos perfectos.

Si bien G_{24} no es un código Golay sino la extensión de uno, debido a la importancia para estudiar a G_{23} consideramos importante incluirlo en este cuadro.

3.3.3. Cotas de Plotkin y Griesmer

Lema 3.3.31 (Desigualdad de Cauchy-Schwarz)**.** *Sea $\{a_1, ..., a_m\}$ y $\{b_1, ..., b_m\}$ conjutos de números reales, con m un entero positivo. Entonces*

$$\left(\sum_{r=1}^{m} a_r b_r\right)^2 \leq \left(\sum_{r=1}^{m} a_r^2\right)\left(\sum_{r=1}^{m} a_r^2\right)$$

Teorema 3.3.32 (Cota de Plotkin)**.** *Sea $q > 1$ un entero. Supongamos que n, d satisfacen que $rn < d$, donde $r = 1 - q^{-1}$. Entonces*

$$A_q(n, d) \leq \left\lfloor \frac{d}{d - rn} \right\rfloor.$$

Demostración. Sea C un código óptimo (n, M, d) sobre el alfabeto A de tamaño q. Sea

$$T = \sum_{c \in C} \sum_{c' \in C} d(c, c')$$

Si $c \neq c'$, con c, $c' \in C$ entonces $d \leq d(c, c')$, y así, como T tiene $M(M-1)$ términos no nulos:

$$M(M-1)d \leq T$$

Consideremos a c y c' como $c = c_1...c_n$ y $c' = c'_1...c'_n$ respectivamente. Sea Z la matriz cuyas filas sean las palabras código de C. Definimos n_{aj} como el número de veces que el elemento $a \in A$ ocurre en la j-ésima columna de Z. Notemos que $\sum_{a \in A} n_{aj} = M$. También definimos α_{aj} como el conjunto de pares ordenados (c, c') tal que c y c' coinciden en la j-ésima coordenada. Luego para un j fijo y $a \in A$ se tiene que

$$|\alpha_{aj}| = \sum_{a \in A} n_{aj}^2.$$

Sabemos que el total de pares ordenados que pueden armarse con las M palabras código es M^2. Así $M^2 - \sum_{a\in A} n_{aj}^2$ representa el número de pares ordenados (c, c') tales que c, c' difieren en la coordenada j-ésima i.e. es $\sum_{c\in C}\sum_{c'\in C} d(c_j, c'_j)$ con j fijo. Luego

$$T = \sum_{c\in C}\sum_{c'\in C} d(c, c') = \sum_{j=1}^{n}\left(\sum_{c\in C}\sum_{c'\in C} d(c_j, c'_j)\right) = \sum_{j=1}^{n}\left(M^2 - \sum_{a\in A} n_{aj}^2\right)$$

Ahora

$$\sum_{j=1}^{n}\left(M^2 - \sum_{a\in A} n_{aj}^2\right) = nM^2 - \sum_{j=1}^{n}\sum_{a\in A} n_{aj}^2.$$

Por la desigualdad de Cauchy-Schwarz con $m = q$ y $a_i = 1$ para todo $i = 1, ..., q$ se tiene que

$$\left(\sum_{a\in A} n_{aj}\right)^2 \leq q\sum_{a\in A} n_{aj}^2$$

y así

$$T = nM^2 - \sum_{j=1}^{n}\sum_{a\in A} n_{aj}^2 \leq nM^2 - \sum_{j=1}^{n} q^{-1}\left(\sum_{a\in A} n_{aj}\right)^2 = nrM^2$$

Finalmente

$$\begin{aligned} M(M-1)d &\leq T \leq M^2 rn \\ M(M-1)d &\leq M^2 rn \\ M &\leq \frac{d}{d-rn} \\ M &\leq \left\lfloor \frac{d}{d-rn} \right\rfloor \end{aligned}$$

□

Existe una versión de la cota de Plotkin para códigos binarios, que da mejores resultados en cuanto a la aproximación de $A_2(n, d)$ condicionado por la paridad de la distancia mínima del código [11],[3] y [4]. Previo a enunciar esta versión, resultará útil para su desarrollo el siguiente lema.

Observación 3.3.33. Sea A un conjunto con q elementos. Si formamos un vector de longitud l con los elementos de A (pudiendose cada elemento repetir), habrá al menos $\frac{l}{q}$ coordenadas de este vector con con algún elemento a de A en todas ellas. Definamos n_a la cantidad de veces que a aparece en el vector formado. Es claro que $\sum_{a\in A} n_a = l$. Supongamos que para cada $a \in A$ se cumple que $n_a < \frac{l}{q}$, es decir, no hay al menos $\frac{l}{q}$ coordenadas de este vector con con algún elemento a de A en todas ellas. Esto no puede ocurrir puesto que que si fuese así $\sum_{a\in A} n_a < l$ [3].

Lema 3.3.34. $A_q(n,d) \leq qA_q(n-1,d)$.

Demostración. Sea C un (n,M,d) código óptimo sobre el alfabeto A de q elementos, i.e. $M = A_q(n,d)$. Definamos C_a al conjunto de palabras códigos pertenecientes a C cuya última coordenada sea a, con $a \in A$. Notemos que existen a lo sumo q conjuntos C_a, con $a \in A$ y que C_a, para al menos un a en A, contiene al menos M/q palabras códigos de C. Suprimamos la última coordenada de este código C_a que contiene al menos M/q palabras código, formando así un nuevo código que llamaremos C_1, un $(n-1,M',d')$ código, con $M' \geq M/q$ y $d \leq d'$.
Si $d' = d$, el lema queda resuelto puesto que $A_q(n,d) \leq qM' \leq qA_q(n-1,d)$. Si $d < d'$ se hace lo siguiente: encontramos las palabras código c, $c' \in C_1$ tales que $d(c,c') = d'$. Se elige una de las coordenadas en donde c y c' difieren y sustituimos en esa coordenada el valor de cero en todas las palabras código de C_1 formando así el código C_2, un $(n-1,M',d'')$ con $d'' \leq d'$. Si $d'' \neq d$ entonces se repiten los pasos anteriores hasta lograr la igualdad con d para concluir así la tésis. □

Teorema 3.3.35 (Cota de Plotkin para códigos binarios)**.** *(a) Si d es par*

$$A_2(n,d) = \begin{cases} 2\left\lfloor \dfrac{d}{2d-n} \right\rfloor & \text{para } n < 2d \\ 4d & \text{para } n = 2d \end{cases}$$

(b) Si d es impar

$$A_2(n,d) = \begin{cases} 2\left\lfloor \dfrac{(d+1)}{(2d+1-n)} \right\rfloor & \text{para } n < 2d+1 \\ 4d+4 & \text{para } n = 2d+1 \end{cases}$$

Demostración. (a) De la demostación del Teorema (3.3.32), siguiendo con la notación usada en el mismo, tenemos que

$$M(M-1)d \leq T \text{ y } T = \textstyle\sum_{j=1}^{n} \left(M^2 - \sum_{a \in A} n_{aj}^2 \right)$$

Observemos que además

$$\sum_{j=1}^{n} \sum_{a \in A} (n_{aj}(M - n_{aj})) = \sum_{j=1}^{n} \left(M^2 - \sum_{a \in A} n_{aj}^2 \right)$$

y como estamos trabajando en el cuerpo binario concluimos que

$$\begin{aligned} T &= \sum_{j=1}^{n} \sum_{a \in A} (n_{aj}(M - n_{aj})) \\ &= \sum_{j=1}^{n} (n_{0j}(M - n_{0j}) + n_{1j}(M - n_{1j})) \\ &= \sum_{j=1}^{n} (2n_{0j}n_{1j}) \end{aligned}$$

Se presentan dos casos: cuando M es par y cuando M es impar.
Cuando M es par, $\sum_{j=1}^{n}(2n_{0j}n_{1j})$ se maximiza cuando $n_{0j} = n_{1j} = \frac{M}{2}$. Luego

$$T \leq \sum_{j=1}^{n}(\frac{M^2}{2}) = n\frac{M^2}{2}$$

y así

$$\begin{aligned} M(M-1)d &\leq n\frac{M^2}{2} \\ dM - d &\leq n\frac{M}{2} \\ 2dM - nM &\leq 2d \\ M &\leq 2\frac{d}{2d-n} \\ M &\leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor \end{aligned}$$

Cuando M es impar, como $n_{0j} + n_{1j} = M$, $\sum_{j=1}^{n}(2n_{0j}n_{1j})$ no se maximiza cuando $n_{0j} = n_{1j} = \frac{M}{2}$, lo hace cuando

$$n_{0j} = \tfrac{M}{2} - \tfrac{1}{2} = \tfrac{M-1}{2} \text{ y } n_{1j} = \tfrac{M}{2} + \tfrac{1}{2} = \tfrac{M+1}{2}$$

Luego $T \leq \sum_{j=1}^{n}\left(2\frac{(M-1)}{2}\frac{(M+1)}{2}\right) = \frac{n}{2}(M-1)(M+1)$ y así

$$\begin{aligned} M(M-1)d &\leq \frac{n}{2}(M-1)(M+1) \\ 2Md &\leq nM + n \\ M\frac{n}{2d-n} &= \frac{2d}{2d-n} - 1. \end{aligned}$$

Como $\lfloor 2x \rfloor \leq 2\lfloor x \rfloor + 1$ es verdadero para cualquier $x \in \mathbb{R}$, se tienen que

$$M \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor - 1 \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor$$

Así como se cumple para M par y M impar, concluimos que esta primera desigualdad queda demostrada.

Para la segunda desigualdad, por el lema anterior, tenemos que

$$A_2(n,d) = A_2(2d,d) \leq 2A_2(2d-1,d)$$

Por la primera desigualdad

$$2A_2(2d-1,d) \leq 2\left(\frac{2d}{2d-(2d-1)}\right) = 4d$$

y así para d par y $n = 2d$

$$A_2(n,d) \leq 4d$$

(b) Sabemos que si d es impar, $A_2(n,d) = A_2(n+1,d+1)$ por Teorema (3.1.9). Como $d+1$ es par y $n+1 < 2d+2$, usamos la primera desigualdad del item anterior para obtener

$$A_2(n+1,d+1) \leq 2\left\lfloor \frac{d+1}{2d+1-n} \right\rfloor.$$

Para la segunda desigualdad consideremos $A_2(2d+1,d)$. Usando los mismos resultados que en la desigualdad anterior se tiene que

$$A_2(2d+1,d) = A_2(2d+2,d+1) \leq 4d+4.$$

□

Teorema 3.3.36 (Cota de Singleton)**.** *Para cualquier entero $q > 1$, cualquier entero positivo n y d tal que $1 \leq d \leq n$ se cumple que*

$$A_q(n,d) \leq q^{n-d+1}.$$

Demostración. Consideremos un (n,M,d) código óptimo C sobre el alfabeto A de q elementos. Sea el código C^* aquel que resulta de haber suprimido las últimas $d-1$ coordenadas en las palabras código de C. Así C^* tiene M palabras código distintas de longitud $n-d+1$ ya que si luego de suprimir las $d-1$ coordenadas de las palabras código de C hubieran resultado dos palabras códigos x^* y y^* iguales, entonces existen las palabras código x y y de C tales que $d(x,y) \leq d-1$, donde x^* y y^* son resultado de suprimir las ultimas $d-1$ coordenadas de x y y. Esto resulta absurdo pues la distancia mínima de C es d.
La cantidad total de elementos de longitud $n-d+1$ sobre el alfabeto A que hay es q^{n-d+1}. Así

$$A_q(n,d) = M \leq q^{n-d+1}.$$

□

Corolario 3.3.37. *Para un $[n,k,d]$ código C sobre el cuerpo F_q se verifica que*

$$k+d \leq n+1$$

Demostración. Por Teorema (3.1.3) sabemos que $B_q(n,d) \leq A_q(n,d)$. De esto y por el teorema anterior se sigue que $q^k \leq A_q(n,d) \leq q^{k-d+1}$ de donde resulta la tésis. □

La siguiente cota que presentaremos es la cota de Griesmer, sólo se aplica a los códigos lineales y es una generalización de la cota de Singleton en el sentido que la cota de Griesmer implica la cota de Singleton para códigos lineales.
Antes de enunciarla, nos resultará útil la definición de código residual que presentamos a continuación.

Definición 3.3.38. Sea C un código lineal de longitud n y $c \in C$ tal que $wt(c) = w$. El *código residual* de C respecto a c, denotado por $\text{Res}(C, c)$, es el código de longitud $n - w$ que se obtiene de C suprimiendo las w coordenadas donde c tiene elementos no nulos.

Lema 3.3.39. *Sea C un código $[n, k, d]$ sobre F_q y $c \in C$ tal que $wt(c) = d$, entonces $Res(C, c)$ es un código $[n - d, k - 1, d']$, donde $d' \geq \left\lceil \frac{d}{q} \right\rceil$*

Demostración. Sin pérdida de generalidad, podemos encontrar un código equivalente a C tal que c tenga unos en las primeras d coordenadas y ceros en el resto.
Por definición $\text{Res}(C, c)$ tiene longitud $n-d$. En la demostración del Teorema (3.1.8) vimos que un código lineal luego de ser perforado da como resultado otro código lineal. De esta forma concluimos que $\text{Res}(C, c)$ es lineal. Pasemos a demostrar que $\dim(\text{Res}(C, c)) = k - 1$.
Definamos la transformación lineal T como sigue

$$\begin{array}{rccc} T: & C & \longrightarrow & \text{Res}(C, c) \\ & x & \longmapsto & x' \end{array}$$

donde x' es el resultado de suprimir las primeras de coordenadas de x. Por la forma en que $\text{Res}(C, c)$ fue construido podemos asegurar que T es sobreyectiva.
Notemos que c pertenece a $\ker(T)$ y así $\dim(\ker(T)) \geq 1$. Sabemos que

$$\dim(\ker(T)) + \dim(\text{Im}(T)) = \dim(C) = k$$

por lo tanto $\dim(\text{Res}(C, c))$ es a lo sumo $k - 1$.
Supongamos que $\dim(\text{Res}(C, c)) < k - 1$, i.e. existe un elemento $l = l_1 l_2 ... l_n$ en $\ker(T)$ y por lo tanto en C, tal que no es múltiplo de c y tal que las últimas $n - d$ coordenadas son nulas. Luego, si l_i de l es no nulo para $1 \leq i \leq d$, como C es lineal, $l - l_i c \in C$. Pero $wt(l - l_i c) < d$ lo cual resulta absurdo pues $d(C) = d$. Este absurdo provino de suponer que $\dim(\text{Res}(C, c)) < k-1$, por lo tanto $\dim(\text{Res}(C, c)) = k-1$.
Finalmente probaremos que $d' \geq \left\lceil \frac{d}{q} \right\rceil$. Sea $x' = x_{d+1} ... x_n$ una palabra código no nula de $\text{Res}(C, c)$ que se corresponde a la palabra código $x = x_1 ... x_d x_{d+1} ... x_n$ de C. Por la Observación (3.3.33) existe $a \in F_q$ tal que al menos $\frac{d}{q}$ coordenadas de $x_1 ... x_d$ son igual a a. Así

$$\begin{aligned} & d \leq wt(x - ac) \leq d - \frac{d}{q} + wt(x_{d+1} ... x_n) \\ & \frac{d}{q} \leq wt(x') \end{aligned}$$

Como x es una palabra código no nula de $\text{Res}(C, c)$ arbitraria, concluimos que $d' \geq \left\lceil \frac{d}{q} \right\rceil$. □

Teorema 3.3.40 (Cota de Griesmer)**.** *Sea C un código $[n, k, d]$ sobre F_q, donde $k \geq 1$. Entonces*

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil .$$

La prueba de este teorema la haremos usando el principio de inducción sobre k.

Demostración. Cuando $k = 1$ se tiene que $n \geq d$ lo que es verdadero.
Cuando $k \geq 1$ usaremos como hipótesis inductiva a $\text{Res}(C, c)$ que, de acuerdo al teorema anterior, es un código $[n - d, k - 1, d']$ con $d' \geq \left\lceil \frac{d}{q} \right\rceil$. Entonces por hipótesis tenemos que

$$n - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil .$$

Como $d' \geq \left\lceil \frac{d}{q} \right\rceil$

$$\begin{aligned} n - d &\geq \sum_{i=0}^{k-2} \left\lceil \frac{d}{q^{i+1}} \right\rceil \\ n &\geq d + \left[\left\lceil \frac{d}{q} \right\rceil + ... + \left\lceil \frac{d}{q^{k-1}} \right\rceil \right] \\ n &\geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil . \end{aligned}$$

Queda así demostrado que se cumple la desigualdad del teorema para cualquier código $[n, k, d]$ con $k \geq 1$. □

Notemos que para $i = 0$

$$\left\lceil \frac{d}{q^0} \right\rceil = d$$

y que

$$\left\lceil \frac{d}{q^i} \right\rceil \geq 1$$

con $i = 1, 2, ..., k - 1$. Teniendo en cuenta esto se sigue que

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil \geq d + k - 1$$

es decir, que la Cota de Griesmer implica la cota de Singleton para códigos lineales.

4. Introducción a los códigos cíclicos

Los códigos cíclicos son códigos lineales que cumplen una condición que hace posible una correspondencia entre sus palabras códigos y polinomios. Así podemos estudiar a los códigos cíclicos analizando un conjunto de polinomios que tiene una estructura algebraica, la de un ideal, y mediante sus elementos es como codificaremos y decodificaremos, usando también conceptos vistos en capítulos anteriores como el peso Hamming y el síndrome de una palabra [12][5]. Recomedamos al lector la lectura del Apéndice antes de leer este capítulo, para recordar algunos conceptos y propiedades de los anillos, estructura que prevalece en este capítulo.

4.1. Códigos cíclicos: polinomio generador

Definición 4.1.1. Un *código cíclico* es un código lineal C que cumple que si $c_0c_1...c_{n-1}$ es una palabra código entonces $c_{n-1}c_0c_1...c_{n-2}$ también lo es.

Ejemplo 4.1.2. (a) El código binario $\{0000, 1010, 0101, 1111\}$ es un código cíclico.

(b) El código lineal $\{x_1x_2x_3 \in F_3^3 \mid \sum_{i=1}^{3} x_i \equiv 0 \,(\text{mod } 3)\}$ con $\{120, 111\}$ una de sus bases, es cíclico. Sus palabras código son

$$\{000, 111, 222, 120, 012, 201, 210, 021, 102\}$$

En este capítulo consideraremos la estructura de anillo que tiene el conjunto F_q^n bajo las siguientes operaciones:

$$a_0a_1a_2...a_{n-1} + b_0b_1b_2...b_{n-1} = a_0 + b_0 \quad a_1 + b_1 \quad a_2 + b_2 \quad ... \quad a_{n-1} + b_{n-1}$$

y

$$a_0a_1a_2...a_{n-1} \odot b_0b_1b_2...b_{n-1} = c_0c_1c_2...c_{n-1} \text{ donde } c_i = \sum_{k+j \equiv i \,(\text{mod } n)} a_kb_j,\ 0 \leq i \leq n-1.$$

Para caracterizar a los códigos cíclicos usaremos el anillo $F_q[x]/x^n - 1$ estableciendo un isomorfismo entre los anillos F_q^n y $F_q[x]/x^n - 1$. De esta manera hacemos corresponder a cada palabra de F_q^n un polinomio de $F_q[x]/x^n - 1$ [4]. Llamemos T a este isomorfismo definido como sigue

$$\begin{array}{rccc} T: & F_q^n & \longrightarrow & F_q[x]/x^n - 1 \\ & a_0a_1...a_{n-1} & \longmapsto & a_0 + a_1x + ... + a_{n-1}x^{n-1} \end{array} \tag{4.1.1}$$

Queda como ejercicio probar que T es una transformación lineal. Probemos que es un isomorfismo.
Probaremos que $\ker(T) = \{0\}$ para demostrar que T es inyectiva. Supongamos por el absurdo que existe $a = a_0a_1...a_{n-1} \in \ker(T)$ y que $a \neq 0$. Así existe $a_j \neq 0$ para algún $0 \leq j \leq n-1$. Luego $T(a_0a_1...a_{n-1}) \neq 0$ pues el término a_jx^j es no nulo. Absurdo, pues $a = a_0a_1...a_{n-1} \in \ker(T)$. El absurdo se produjo al suponer que existe $a = a_0a_1...a_{n-1} \in \ker(T)$ y que $a \neq 0$. Luego $\ker(T) = \{0\}$ y por lo tanto T es inyectiva.
Sea $a_0 + a_1x + ... + a_{n-1}x^{n-1} \in F_q[x]/x^n - 1$. Existe un elemento $a \in F_q^n$ tal que $T(a) = a_0 + a_1x + ... + a_{n-1}x^{n-1}$ y es $a_0a_1...a_{n-1}$ y con esto concluimos que T es sobreyectiva. Así T es un isomorfismo.

Notemos que cualquier polinomio de la forma ax^n con $a \in F_q$ es congruente módulo $x^n - 1$ a a, en particular cuando $a = 1$ se tiene que $x^n \equiv 1(\mod x^n - 1)$.
Ahora, dado el polinomio $a(x) = a_0 + a_1x + ... + a_{n-1}x^{n-1} \in F_q[x]/x^n - 1$, operando siempre en $F_q[x]/x^n - 1$, tenemos:

$$\begin{aligned} xa(x) &= a_0x + a_1x^2 + ... + a_{n-1}x^n \\ &= a_{n-1} + a_0x + ... + a_{n-2}x^{n-1} \end{aligned}$$

$$\begin{aligned} x^2a(x) &= a_0x^2 + a_1x^3 + ... + a_{n-2}x^n + a_{n-1}x^{n+1} \\ &= a_{n-2} + a_{n-1}x + a_0x^2 + ... + a_{n-3}x^{n-1} \end{aligned}$$

$$\begin{aligned} x^ia(x) &= a_0x^i + a_1x^{i+1} + ... + a_{n-1}x^{n-1+i} \\ &= a_{n-i} + a_{(n-i)+1}x + ... + a_0x^i + ... + a_{n-(i+1)}x^{n-1} \end{aligned}$$

con $1 \leq i \leq n-1$.
Luego si la imágen de $a_0a_1...a_{n-1}$ por T es $a(x) = a_0 + a_1x + ... + a_{n-1}x^{n-1}$, es decir

$$a_0a_1...a_{n-1} \quad \longmapsto \quad a(x) = a_0 + a_1x + ... + a_{n-1}x^{n-1}$$

se tiene que

$$(a_{n-i}...a_0a_1...a_{n-(i+1)}) \quad \longmapsto \quad x^ia(x) = a_{n-i} + ... + a_0x^i + a_1x^{i+1}... + a_{n-(i+1)}x^{n-1}.$$

Dada una palabra código de un código cíclico pudimos encontrar en $F_q[x]/x^n - 1$ una forma de obtener las palabras código que resultan de mover r veces a la derecha los elementos de cada coordenada de la palabra código dada. Convendremos en ver una palabra código simultáneamente como un elemento de F_q^n y como un polinomio de $F_q[x]/x^n - 1$.

Teorema 4.1.3. *Sea T la transformación lineal definida en (4.1.1). Un subconjunto no vacío C de F_q^n es un código cíclico si y sólo si $T(C)$ es un ideal de $F_q[x]/x^n - 1$.*

Demostración. $\Rightarrow$) Supongamos que C es un código cíclico. Para demostrar que $T(C)$ es un ideal de $F_q[x]/x^n-1$ probaremos que $T(C)$ es no vacío y que para todo $c(x), c'(x), a(x) \in T(C)$ y $g(x) \in F_q[x]/x^n-1$ se cumple que $c(x)-c'(x) \in T(C)$ y $g(x)a(x) \in T(C)$.
Notemos que $T(C)$ es no vacío debido a que C es distinto del vacío. Ahora, sean c y c' dos palabras código de C, pre imágenes por T de $c(x)$ y $c'(x)$ respectivamente. Sabemos que $c-c' \in C$ por ser C un código lineal. Luego se tiene que $T(c-c') \in T(C)$ y como T es una transformación lineal

$$T(c-c') = T(c) - T(c') = c(x) - c'(x) \in T(C).$$

Para demostrar que $g(x)a(x) \in T(C)$ consideremos a $g(x) = g_0+g_1x+...+g_{n-1}x^{n-1}$. Se tiene que:

$$\begin{aligned} g(x)a(x) &= (g_0 + g_1x + ... + g_{n-1}x^{n-1})a(x) \\ &= g_0a(x) + (g_1x)a(x) + ... + (g_{n-1}x^{n-1})a(x) \\ &= g_0a(x) + g_1(xa(x)) + ... + g_{n-1}(x^{n-1}a(x)) \\ &= \sum_{i=0}^{n-1} g_i(x^ia(x)) \end{aligned}$$

y por la observación anterior sabemos que cada término de esta suma pertenece a $T(C)$ por lo tanto $g(x)a(x) \in T(C)$. Con esto concluimos que $T(C)$ es un ideal de $F_q[x]/x^n-1$.

$\Leftarrow$) Supongamos ahora que $T(C)$ es un ideal de $F_q[x]/x^n-1$. Probemos que C es un código lineal. Sean $\alpha \in F_q \subset F_q[x]/x^n-1$ y $a, b \in C$. Se tiene que $\alpha T(a) \in T(C)$ por ser $T(C)$ un ideal y como $T(b) \in T(C)$ se tiene que $\alpha T(a) + T(b) \in T(C)$ por ser $T(C)$ un ideal. Como T es una transformación lineal

$$\begin{aligned} \alpha T(a) + T(b) &= T(\alpha a) + T(b) \\ &= T(\alpha a + b) \end{aligned}$$

es decir, $\alpha a + b \in C$ y así C es un código lineal.

Probemos ahora que C es cíclico. Sea $c = (c_0c_1...c_{n-1})$ una palabra código de C y $T(c) = c_0 + c_1x + ... + c_{n-1}x^{n-1}$. Sabemos que $xT(c) \in T(C)$ ya que $T(C)$ es un ideal, y como

$$\begin{aligned} xT(C) &= x(c_0 + c_1x + ... + c_{n-1}x^{n-1}) \\ &= c_0x + c_1x^2 + ... + c_{n-1}x^n \\ &= c_{n-1} + c_0x + c_1x^2 + ... + c_{n-2}x^{n-1} \end{aligned}$$

se tiene que $c_{n-1}c_0c_1...c_{c-2} \in C$ y por lo tanto C es un código cíclico. $\square$

Ejemplo 4.1.4. Considere el código cíclico binario $C = \{000, 110, 011, 101\}$. El conjunto formado por sus elementos correspondientes en $F_2[x]/x^3 - 1$ por T es

$$\{0, 1 + x, x + x^2, 1 + x^2\}.$$

Sabiendo que

$$F_2[x]/x^3 - 1 = \{0, 1, 1 + x, x + x^2, 1 + x^2, x, x^2, 1 + x + x^2\}.$$

demostremos que $\{0, 1 + x, x + x^2, 1 + x^2\}$ es un ideal. Queda como ejercicio para el lector demostrar que la suma es cerrada en este conjunto.
La siguiente tabla muestra que para $a(x) \in \{0, 1 + x, x + x^2, 1 + x^2\}$ y $g(x) \in F_2[x]/x^3 - 1$ se tiene que $a(x)g(x) \in \{0, 1 + x, x + x^2, 1 + x^2\}$.

$\cdot$	0	1	$1+x$	$x+x^2$	$1+x^2$	x	x^2	$1+x+x^2$
0	0	0	0	0	0	0	0	0
$1+x$	0	$1+x$	$1+x^2$	$1+x$	$x+x^2$	$x+x^2$	$1+x^2$	0
$x+x^2$	0	$x+x^2$	$1+x$	$x+x^2$	$1+x^2$	$1+x^2$	$1+x$	0
$1+x^2$	0	$1+x^2$	$x+x^2$	$1+x^2$	$1+x$	$1+x$	$x+x^2$	0

Teorema 4.1.5. *Sea I un ideal de $F_q[x]/x^n - 1$. Sea $g(x)$ un polinomio mónico de menor grado en I. Entonces*

(a) $g(x)$ es el único polinimio mónico de grado mínimo en I, además, este polinomio genera a I i.e. $I = \langle g(x) \rangle$.

(b) $g(x)|x^n - 1$.

(c) Sea C el código cíclico correspondiente al ideal I. Si $gr(g(x)) = r$, entonces C tiene dimensión k, con $k = n - r$. Más aún

$$I = \langle g(x) \rangle = \{a(x)g(x) : \ gr(a(x)) < n - r\}$$

Demostración. (a) Supongamos que $g_1(x)$ y $g_2(x)$ son dos polinomios mónicos distintos de grado mínimo r. Luego $g_1(x) - g_2(x)$ es un polinomo no nulo de grado menor que r. Si a es el coeficiente principal de $g_1(x) - g_2(x)$, entonces $a^{-1}(g_1(x) - g_2(x))$ es un polinomio mónico de grado menor que r. Absurdo. Este absurdo provino de suponer que había dos polinomios mónicos de menor grado, por lo tanto $g(x)$ es único.

Sea $f(x) \in I$. Por el algoritmo de la división de polinomios se tiene que

$$f(x) = q(x)g(x) + r(x)$$

con $q(x), r(x) \in F_q[x]/x^n - 1$ y $\text{gr}(r(x)) < \text{gr}(g(x))$ o $r(x) = 0$. Como $q(x)g(x) \in I$ debido a que I es un ideal y $r(x) = f(x) - q(x)g(x)$, se tiene que $r(x) \in I$. Si $r(x) \neq 0$ y además es mónico, no puede ocurrir que $r(x) \in I$ ya que

se contradice el hecho $g(x)$ es el único polinimio mónico de grado mínimo en I. Si $r(x) \neq 0$ y no es mónico, entonces $r_j^{-1} r(x) \in I$, con r_j el coeficiente principar de $r(x)$, lo que no puede ocurrir ya que $\text{gr}(r(x)) = \text{gr}(r(x)) < \text{gr}(g(x))$ y $g(x)$ es el único polinimio mónico de grado mínimo en I. Así debe ocurrir que $r(x) = 0$.
Luego $f(x) = q(x)g(x)$ y como $f(x)$ es un polinomio arbitrario de I concluimos que

$$I = \langle g(x) \rangle = \{h(x)g(x) : \, h(x) \in F_q[x]/x^n - 1\}$$

(b) Sabemos que

$$x^n - 1 = s(x)g(x) + r(x)$$

con $\text{gr}(r(x)) < \text{gr}(g(x))$ o $r(x) = 0$. Sabemos que $x^n - 1$ es el elemento neutro de suma en $F_q[x]/x^n - 1$ y, por lo tanto, también en I. Como

$$r(x) = x^n - 1 - s(x)g(x)$$

y $\text{gr}(r(x)) < \text{gr}(g(x))$ o $r(x) = 0$, $r(x) \in I$. Haciendo razonamientos análogos a los realizados en el ítem anterior, concluimos que $r(x) = 0$ ya que $g(x)$ es el polinomio mónico de menor grado en I. Así $g(x) | x^n - 1$.

(c) Demostraremos primero que

$$I = \langle g(x) \rangle = \{a(x)g(x) : \, \text{gr}(a(x)) < n - r\}$$

Sabemos que $I = \langle g(x) \rangle = \{f(x)g(x) : \, f(x) \in F_q[x]/x^n - 1\}$. Lo que haremos a continuación será reducir los polinomios $f(x) \in F_q[x]/x^n - 1$ a polinomios de grado menor a $k = n - r$.
Por inciso anterior $x^n - 1 = h(x)g(x)$, con $\text{gr}(h(x)) = n - r$. Esto quiere decir que $h(x)g(x) \equiv 0 \pmod{x^n - 1}$. Si dividimos $f(x)$ por $h(x)$ tenemos que

$$f(x) = q(x)h(x) + r(x)$$

con $q(x) \in F_q[x]/x^n - 1$ y $\text{gr}(r(x)) < k = n - r$ o $r(x) = 0$. Multiplicando miembro a miembro por $g(x)$

$$\begin{aligned} f(x)g(x) &= q(x)h(x)g(x) + r(x)g(x) \\ f(x)g(x) &= q(x)0 + r(x)g(x) \\ f(x)g(x) &= r(x)g(x) \end{aligned}$$

De esta última igualdad se deduce que

$$I = \langle g(x) \rangle = \{a(x)g(x) : \, \text{gr}(a(x)) < n - r\}$$

y como la cantidad de polinomios $a(x)$ que se pueden formar, con $\text{gr}(a(x)) < k = n - r$ o $a(x) = 0$, es q^k, teniendo en cuenta la Observación (2.1.8), concluimos que I, y por lo tanto también C, tiene dimensión k [13].

□

Definición 4.1.6. El único polinomio mónico de menor grado de un ideal I de $F_q[x]/x^n-1$ es llamado *polinomio generador* de I. Sea C un código cíclico, el polinomio generador de $T(C)$ es también llamado *polinomio generador* de C.

Ejemplo 4.1.7. (a) Consideremos el código cíclico binario $C = \{000, 110, 011, 101\}$ y su correspondiente ideal $\{0, 1+x, x+x^2, 1+x^2\} \subset F_2[x]/x^3-1$.
Observemos que el polinomio generador del ideal es $1+x$, que $1+x \mid x^3-1$, que $\mathrm{gr}(1+x) = 1$ por lo tanto $\dim(C) = 2$. Los polinomios de grado menor a 2 en $F_2[x]/x^3-1$ son $0, 1, x, x+1$, y ayudados con la tabla del ejemplo anterior podemos verificar que

$$\{0, 1+x, x+x^2, 1+x^2\} = \{0(1+x), 1(1+x), x(1+x), x+1(1+x)\}$$

(b) Al código cíclico binario $\{0000, 1010, 0101, 1111\}$ le corresponde el ideal $\{0, 1+x^2, x+x^3, 1+x+x^2+x^3\} \subset F_2[x]/x^4-1$.
Nuevamente observemos que $1+x^2$ es el polinomio generador del código cíclico, que es de grado 2, y que el código tiene dimensión 2. Sabiendo que los polinomios de grado menor a 2 en $F_2[x]/x^4-1$ son $0, 1, x, x+1$ y teniendo en cuenta que

$\cdot$	0	1	x	$x+1$
$1+x^2$	0	$1+x^2$	$x+x^3$	$1+x+x^2+x^3$

verificamos que

$$\{0, 1+x^2, x+x^3, 1+x+x^2+x^3\} = \{f(x)(1+x^2) : \mathrm{gr}(f(x)) < 2\}$$

Observación 4.1.8. Notemos que para cualquier $f(x) \in F_q[x]/x^n-1$, $\langle f(x)\rangle$ resulta un ideal de $F_q[x]/x^n-1$. En efecto, es claro que es distinto del vacío y que para cualquier $a(x), b(x) \in \langle f(x)\rangle$ $(a(x)-b(x)) \in \langle f(x)\rangle$. Ahora, sea $i(x) \in \langle f(x)\rangle$ y $r(x) \in F_q[x]/x^n-1$, se cumple que

$$i(x)r(x) = (a(x)f(x))r(x) = (a(x)r(x))f(x) \in \langle f(x)\rangle$$

con $a(x) \in F_q[x]/x^n-1$. Esto quiere decirnos que el polinomio generador de un ideal puede que no sea el único polinomio que genere ese ideal.

Teorema 4.1.9. *Cada divisor mónico de x^n-1 es el polinomio generador de algún código cíclico en F_q^n.*

Demostración. Sea $g(x)$ un divisor mónico de x^n-1 y I el ideal $\langle g(x)\rangle$ de $F_q[x]/x^n-1$. Sea $h(x)$ el polinomio generador de I. Como $g(x)$ genera a I y $h(x) \in I$, se tiene que $h(x) = a(x)g(x)$ (más bien $h(x) \equiv a(x)g(x) \pmod{x^n-1}$) para algún $a(x) \in F_q[x]/x^n-1$ y así $g(x)|h(x)$. Por otro lado, como $h(x)$ genera a I, se tiene que $g(x) = b(x)h(x)$ para algún $a(x) \in F_q[x]/x^n-1$ (más bien $g(x) \equiv b(x)h(x) \bmod(x^n-1)$) y así $h(x)|g(x)$. Al ser $g(x)$ y $h(x)$ mónicos, concluimos que $g(x) = h(x)$ y así $g(x)$ es el polinomio generador de algún código cíclico en F_q^n. □

Ejemplo 4.1.10. Consideremos $x^4 - 1$ sobre el cuerpo F_3. Éste queda factorizado como

$$x^4 - 1 = (x-1)(x^3 + x^2 + x + 1) = (x-1)(x-2)(x^2+1),$$

así los códigos cíclicos de longitud 4 sobre F_3 son 8 y tienen como polinomios generadores a uno de los siguientes polinomios

$$1 \qquad x-1 \qquad x-2 \qquad x^2+1$$

$$(x-1)(x-2) \quad (x-1)(x^2+1) \quad (x-2)(x^2+1) \quad (x-1)(x-2)(x^2+1)$$

Nota 4.1.11. El lector ya habrá notado que el código cíclico con polinomio generador $(x-1)(x-2)(x^2+1) = x^4 - 1$ es $\{0000\}$ y que el código cíclico con polinomio generador 1 es F_q^n.

4.2. Matriz generadora y matriz de control de paridad

En el Capitulo 2 concluimos que un código lineal está completamente determinado por alguna matriz generadora del código o por alguna matriz de control de paridad. Presentamos una forma de ambas llamadas estándard, concluyendo que eran convenientes para facilitar el uso y estudio de las propiedades de un código lineal. Un código cíclico es un código lineal, por lo que también posee matrices generadoras y de control de paridad. En esta sección presentamos una forma que posee la matriz generadora y la matriz de control de paridad de un código cíclico, las cuales están construidas a partir de una base especial determinada a partir del isomorfismo T definido en la sección anterior.

Teorema 4.2.1. *Sea $g(x) = g_0 + g_1 x + ... + g_r x^r$ el polinomio generador de un código cíclico de longitud n. Entonces $g_0 \neq 0$.*

Demostración. Sea C un código cíclico de longitud n con polinomio generador $g(x) = g_0 + g_1 x + ... + g_r x^r$, i.e. $g_r = 0$, y sea I el ideal que se corresponde a C por T.
Supongamos que $g_0 = 0$. Sabemos que $x^{n-1} \in F_q[x]/x^n - 1$, luego $x^{n-1} g(x) \in I$. Ya que $x^n \equiv 1(\text{ mod } x^n - 1)$ se tiene que

$$x^{n-1} g(x) = \frac{1}{x} g(x) = g_1 + g_2 x + ... + g_r x^{r-1} \in I$$

esto es absurdo ya que $g(x)$ es el polinomio generador i.e el polinomio mónico de menor grado en I. El absurdo provino de suponer que $g_0 = 0$, por lo que concluimos que $g_0 \neq 0$. □

Teorema 4.2.2. *Sea C un código cíclico de longitud n con polinomio generador $g(x) = g_0 + g_1x + ... + g_rx^r$. Entonces una matriz generadora de tamaño $(n-r) \times n$ del código C es*

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \cdots & g_r & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_r \end{pmatrix}$$

Demostración. Sea I el ideal que se corresponde a C por T. Nuestra prueba consistirá en entontrar una base para I y así una base para C.
Supongamos que la longitud de C es n. Por Teorema (4.1.5) sabemos que $\dim(C) = \dim(I) = k = n - r$ y que
$I = \{f(x)g(x) : \text{gr}(f(x)) < k = n-r\}$, además, como $g(x)$ es el polinomio generador de C, éste es mónico, i.e. $g_r = 1$. Consideremos el conjunto $\{g(x), xg(x), x^2g(x), ..., x^{k-1}g(x)\}$. Probaremos que este conjunto es linealmente independiente.

$$a_0g(x) + a_1xg(x) + a_2x^2g(x) + ... + a_{k-1}x^{k-1}g(x) = 0 \qquad a_i \in F_q,\ i = 0, ..., k-1$$

desarrollando el miembro de la izquierda se tiene que

$$a_0g_0+(a_0g_1+a_1g_0)x+(a_0g_2+a_1g_1+a_2g_0)x^2+...+\left(\sum_{i+j=k-2} a_ig_j\right)x^{r+k-2}+a_{k-1}g_rx^{r+k-1} = 0$$

Ahora, $a_0g_0 = 0$, $a_0g_1 + a_1g_0 = 0$,... $\sum_{i+j=k-2} a_ig_j = 0$ y $a_{k-1}g_r = 0$ ya que el conjunto $\{1, x, x^2, ..., x^{r+k-1}\}$=$\{1, x, x^2, ..., x^{n-1}\}$ es linealmente independiente.
Por el teorema anterior $g_0 \neq 0$, por lo tanto de $a_0g_0 = 0$ se tiene que $a_0 = 0$. Luego de $a_0g_1 + a_1g_0 = 0$, como $a_0 = 0$, se tiene que $a_1 = 0$. Y continuando de la misma forma se tiene que $a_i = 0,$ con $i = 0, ..., k-1$. Esto quiere decir que
$\{g(x), xg(x), x^2g(x), ..., x^{k-1}g(x)\}$ que posee $k = n - r$ elementos, es linealmente independiente y por lo tanto una base de I.
Sabemos que

$$\begin{array}{ccc} & T^{-1} & \\ g(x) & \longmapsto & g_0g_1...g_r0..,0 \\ xg(x) & \longmapsto & 0g_0g_1...g_r0..,0 \\ x^2g(x) & \longmapsto & 00g_0g_1...g_r0..,0 \\ & \vdots & \\ x^{k-1}g(x) & \longmapsto & 0..,0g_0g_1...g_r \end{array} \tag{4.2.1}$$

por lo tanto la matriz generatriz de C es

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \cdots & g_r & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_r \end{pmatrix}$$

□

Ejemplo 4.2.3. (a) El código cíclico binario $\{000, 110, 011, 101\}$ tiene por polinomio generador a $1 + x$. Una matriz generadora de este código es

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

(b) El código binario cíclico $\{0000, 1010, 0101, 1111\}$ tiene por polinomio generador a $1 + x^2$. Así una matriz generadora de este código es

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

(c) x^5-1 sobre el cuerpo binario queda factorizado como $(x+1)(x^4+x^3+x^2+x+1)$. Las matrices generadoras de los códigos cíclicos asociados a cada factor de $x^5 - 1$ son

Polinomio generador	Matriz generadora
1	$\begin{pmatrix} I_5 \end{pmatrix}$
$x + 1$	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$
$x^4 + x^3 + x^2 + x + 1$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$
$x^5 - 1$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

(d) $x^4 - 1$ sobre el cuerpo F_3 queda factorizado como

$$x^4-1 = (x-1)(x^3+x^2+x+1) = (x-1)(x-2)(x^2+1) = (x+2)(x+1)(x^2+1).$$

Las matrices generadoras de los códigos cíclicos asociados a cada factor de

$x^4 - 1$ son

Polinomio generador	Matriz generadora
1	$\begin{pmatrix} I_4 \end{pmatrix}$
$x+1$	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
$x+2$	$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}$
x^2+1	$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$
$(x+1)(x+2)$	$\begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}$
$(x+1)(x^2+1)$	$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$
$(x+2)(x^2+1) = x^3+x^2+x+1$	$\begin{pmatrix} 1 & 2 & 1 & 2 \end{pmatrix}$
$x^5 - 1 = 0$	$\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$

Dado un código cíclico C es natural preguntarse si $C^{\perp}$ también es un código cíclico. La respuesta es sí, $C^{\perp}$ es un código cíclico. Para poder demostrar esto definimos la función σ de la siguiente forma

$$\begin{array}{rccc} \sigma: & F_q^n & \longrightarrow & F_q^n \\ & a_0a_1...a_{n-1} & \longmapsto & a_{n-1}a_0a_1...a_{n-2} \end{array} \tag{4.2.2}$$

Así σ es la función que hace que los valores de las coordenadas de una palabra de F_q^n se muevan una posición a la derecha. Componiendo σ k veces, podemos obtener k movimientos a la derecha en una palabra dada.
La composición de k veces σ aplicado a un elemento $a \in F_q^n$ lo denotaremos como $\sigma^k(a)$ y definimos a $\sigma^0(a) = Id(a) = a$, donde Id es la función identidad.
El lector habrá notado que la cantidad de movimientos que se deben realizar a derecha para volver a obtener la palabra dada al comienzo es de n. Esto nos quiere decir que para los valores de los supraíndices podemos encontrar otros equivalente módulo n que nos da exactamente la misma cantidad de movimientos a derecha.

Sean $a, b \in F_q^n$ tales que $a = a_0a_1...a_{n-1}$ y $b = b_0b_1...b_{n-1}$. Notemos que para $i \geq 0$

$$\sigma(a) \cdot \sigma^i(b) = a \cdot \sigma^{i-1}(b) \text{ donde } \cdot \text{ denota el producto escalar.}$$

Demostremos esto por inducción sobre i.

Para $i = 0$ tenemos que

$$\begin{aligned}\sigma(a) \cdot \sigma^0(b) &= \sigma(a) \cdot b \\ &= (a_{n-1}a_0a_1...a_{n-2}) \cdot (b_0b_1b_2...b_{n-1}) \\ &= a_{n-1}b_0 + a_0b_1 + a_1b_2 + ... + a_{n-2}b_{n-1} \\ &= a_0b_1 + a_1b_2 + ... + a_{n-2}b_{n-1} + a_{n-1}b_0 \\ &= a \cdot \sigma^{n-1}(b)\end{aligned}$$

y como $-1 \equiv n - 1 \pmod{n}$ concluimos que la igualdad a demostrar es verdadera para $i = 0$.
Supongamos ahora que $\sigma(a) \cdot \sigma^k(b) = a \cdot \sigma^{k-1}(b)$ es verdadero.

$$\begin{aligned}\sigma(a) \cdot \sigma^{k+1}(b) &= \sigma(a) \cdot \sigma^k(\sigma(b)) \\ &= a \cdot \sigma^{k-1}(\sigma(b)) \\ &= a \cdot \sigma^k(b)\end{aligned}$$

Con esto concluimos que $\sigma(a) \cdot \sigma^i(b) = a \cdot \sigma^{i-1}(b)$ es verdadero para todo entero $i \geq 0$.

Teorema 4.2.4. *El código dual de un código cíclico es cíclico.*

Demostración. Sea C un código cíclico. Sabemos que $C^\perp$ es un código lineal, solo basta demostrar que si $h \in C^\perp$, entonces $\sigma(h) \in C^\perp$, con σ definida en (4.2.2).
Sea $c \in C$, entonces

$$\sigma(h) \cdot c = \sigma(h) \cdot \sigma^n(c) = h \cdot \sigma^{n-1}(c) = 0$$

ya que $\sigma^{n-1}(c) \in C$. Así $\sigma(h) \in C^\perp$. □

Al ser el código dual de un código cíclico un código cíclico, lo que sigue es determinar el polinomio generador del mismo.

Teorema 4.2.5. *Sea $C \subset F_q^n$ un código cíclico y $c \in C$. Sea $g(x)$ el polinomio generador de C y $h(x)$ un polinomio tal que $x^n - 1 = g(x)h(x)$. Entonces $c \in C$ si y sólo si $c(x)h(x) = 0$ (en $F_q[x]/x^n - 1$), donde $c(x)$ es el correspondiente de c por T.*

Demostración. $\Rightarrow$) Ya que $g(x)$ es la generatriz del código, para algún $a(x) \in F_q[x]/x^n - 1$, se cumple que $c(x) = a(x)g(x)$. Así, teniendo en cuenta que $x^n - 1 = g(x)h(x) = 0$ (en $F_q[x]/x^n - 1$), se tiene que

$$\begin{aligned}c(x)h(x) &= a(x)g(x)h(x) \\ &= a(x) \cdot 0 \\ &= 0\end{aligned}$$

$\Leftarrow$) Sea $c(x)h(x) = 0$. Por el algoritmo de la división $c(x) = q(x)g(x) + r(x)$ con $q(x), r(x) \in F_q[x]/x^n - 1$ y $\text{gr}(r(x)) < \text{gr}(g(x))$ o $r(x) = 0$. Supongamos por el absurdo que $\text{gr}(r(x)) < \text{gr}(g(x))$.
Multiplicando miembro a miembro por $h(x)$ y teniendo en cuenta que $x^n - 1 = g(x)h(x) = 0$ (en $F_q[x]/x^n - 1$), se tiene que

$$\begin{aligned} c(x)h(x) &= q(x)g(x)h(x) + r(x)h(x) \\ 0 &= q(x)(x^n - 1) + r(x)h(x) \\ 0 &= r(x)h(x) \end{aligned}$$

y así $r(x)$ pertenece al ideal correspondiente a C por T, con $\text{gr}(r(x)) < \text{gr}(g(x))$. Absurdo. Luego $r(x) = 0$, $c(x) = q(x)g(x)$ y por lo tanto $c \in C$. □

Consideremos C, $g(x)$ y $h(x)$ como en el teorema anterior y $\text{gr}(g(x)) = r$. Así $\text{gr}(h(x)) < n - r = k$ y por lo tanto $\dim(\langle h(x) \rangle) = n - k = \dim(C^\perp)$. Podemos pensar que, si $h(x)$ es el correspondiente a la palabra h por T, $h \in C^\perp$ y $h(x)$ es el polinomio generador de $C^\perp$. Ambas cosas no siempre son verdaderas.

Ejemplo 4.2.6. (a) $x^4 - 1$ sobre el cuerpo F_3 queda factorizado como

$$\begin{aligned} x^4 - 1 &= (x-1)(x^3 + x^2 + x + 1) = (x-1)(x-2)(x^2+1) \\ &= (x+2)(x+1)(x^2+1) \\ &= (x^2+2)(x^2+1). \end{aligned}$$

El código cíclico sobre F_3 con polinomio generador $x^2 + 2$ tiene como una matriz generadora a

$$G_1 = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}$$

Suponiendo que $(x^2 + 1)$ es el polinomio generador del código dual de este código cíclico, tenemos que una matriz de control de paridad es

$$H_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

y como

$$G_1 H_1^\top = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

podemos concluir que $x^2 + 1$ es ciertamente el polinomio generador del código dual del código cíclico generado por $x^2 + 2$.

(b) $x^7 - 1$ en el cuerpo binario queda factorizado como

$$\begin{aligned} x^7 - 1 &= (x-1)(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1) \\ &= (x+1)(x^3 + x + 1)(x^3 + x^2 + 1) \\ &= (x^4 + x^3 + x^2 + 1)(x^3 + x^2 + 1) \end{aligned}$$

El código cíclico binario generado por $(x^4+x^3+x^2+1)$ tiene como una matriz generadora a

$$G_2 = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Suponiendo que $x^3 + x^2 + 1$ es el polinomio generador del código dual de este código cíclico, tenemos que una matriz de control de paridad es

$$H_2 = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

sin embargo podemos observar que

$$G_2 H_2^\top = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Luego $x^3 + x^2 + 1$ no es el polinomio generador del código dual del generado por $x^4 + x^3 + x^2 + 1$ y si $x^3 + x^2 + 1$ es el correspondiente de h por T, h no pertenece a este código dual.

Definición 4.2.7. Sea $h(x) = \sum_{i=0}^{k} a_i x^i$ un polinomio de grado k ($a_k \neq 0$) sobre F_q. El *polinomio recíproco* $h_R(x)$ de $h(x)$ está definido por

$$h_R(x) = x^k h\left(\frac{1}{x}\right) = \sum_{i=0}^{k} a_{k-i} x^i$$

Ejemplo 4.2.8. (a) Sea $h(x) = 2+2x^2+x^3+2x^6 \in F_3[x]$. El polinomio recíproco de $h(x)$ es

$$\begin{aligned} h_R(x) &= x^6 h\left(\frac{1}{x}\right) \\ &= x^6 \left(2 + \frac{2}{x^2} + \frac{1}{x^3} + \frac{2}{x^6}\right) \\ &= 2 + x^3 + 2x^4 + 2x^6 \end{aligned}$$

(b) En el último ejemplo dijimos que x^7-1 en el cuerpo binario queda factorizado como

$$\begin{aligned} x^7-1 &= (x-1)(x^6+x^5+x^4+x^3+x^2+x+1) \\ &= (x+1)(x^3+x+1)(x^3+x^2+1) \end{aligned}$$

Notemos que el polinomio recíproco de x^3+x+1 es

$$x^3\left(1+\frac{1}{x}+\frac{1}{x^3}\right) = x^3+x^2+1$$

que es otro factor de x^7-1

(c) x^4-1 sobre el cuerpo F_3 queda factorizado como

$$\begin{aligned} x^4-1 &= (x-1)(x^3+x^2+x+1) = (x-1)(x-2)(x^2+1) \\ &= (x+2)(x+1)(x^2+1) \\ &= (x^2+2)(x^2+1). \end{aligned}$$

el polinomio recíproco de x^2+2 es

$$x^2\left(2+\frac{1}{x^2}\right) = 2x^2+1$$

y $2x^2+1|x^4-1$ ya que $x^4-1 = 2(x^2+1)(2x^2+1)$.
El lector puede verificar que el polinomio recíproco de x^2+1 es el mismo.

Observación 4.2.9. 1. Del último ejemplo, $x^4-1 = (x^2+2)(x^2+1)$ sobre el cuerpo ternario. La palabra c quien le corresponde x^2+2 por T es 2010. Estamos tentados a decir que la palabra a quien le corresponde el recíproco de x^2+2 es 0102, pero esto, como se puede verificar, es falso. Sin embargo, si hacemos tres movimientos a derecha en los valores de cada posición de 0102 obtenemos 1020 que si es la palabra a la que se corresponde el polinomio reciproco de x^2+2. En general, si $\text{gr}(h(x)) = k$, se tiene que la palabra correspondiente al polinomio recíproco de $h(x)$ en F_q^n se obtiene haciendo $k+1$ movimientos a derecha en los valores de cada posición de la palabra $h_{n-1}h_{n-2}...h_1h_0$.

2. Si $h(x)|x^n-1$ entonces $h_R(x)|x^n-1$.
Supongamos que $\text{gr}(h(x)) = k$. Hagamos $x^n-1 = g(x)h(x)$, con $\text{gr}(g(x)) = n-k$. Entonces se tiene que $x^{-n}-1 = g(x^{-1})h(x^{-1})$. Multiplicando por x^n miembro a miembro se tiene que

$$\begin{aligned} x^n(x^{-n}-1) &= x^n g(x^{-1})h(x^{-1}) \\ 1-x^n &= x^{n-k}g(x^{-1})x^k h(x^{-1}) \\ 1-x^n &= g_R(x)h_R(x) \\ x^n-1 &= -g_R(x)h_R(x). \end{aligned}$$

Así $h_R(x)|x^n-1$.

Teorema 4.2.10. *Sea $g(x)$ el polinomio generador del $[n,k]$ código C sobre F_q^n. Sea $h(x) = (x^n-1)/g(x)$. Entonces $h_0^{-1}h_R(x)$ es el polinomio generador de $C^\perp$, donde h_0 es el término constante de $h(x)$.*

Demostración. Hagamos $g(x) = \sum_{i=0}^{n-1} g_i x^i$ y $h(x) \sum_{i=0}^{n-1} h_i x^i$. Sea $k = \mathrm{gr}(h(x))$.

$$\begin{aligned}
0 \equiv & g(x)h(x) \\
\equiv & (g_0h_0 + g_0h_1x + ... + g_0h_{n-1}x^{n-1}) + (g_1h_0x + g_1h_1x^2 + ... + g_1h_{n-2}x^{n-1} \\
& + g_1h_{n-1}) + (g_2h_0x^2 + ... + g_2h_{n-3}x^{n-1} + g_2h_{n-2} + g_2h_{n-1}x) + ... \\
& + (g_{n-2}h_0x^{n-2} + g_{n-2}h_1x^{n-1} + g_{n-2}h_2 + g_{n-2}h_3x + ... + g_{n-2}h_{n-1}x^{n-3}) \\
& + (g_{n-1}h_0x^{n-1} + g_{n-1}h_1 + g_{n-1}h_2x + g_{n-1}h_3x^2 \\
& + ... + g_{n-1}h_{n-2}x^{n-1} + g_{n-1}h_{n-1}x^{n-2}) \\
\equiv & (g_1h_{n-1} + g_2h_{n-2} + .. + g_{n-1}h_1 + g_0h_0) + (g_2h_{n-1} + g_3h_{n-2} + .. + g_0h_1 \\
& + g_1h_0)x + (g_3h_{n-1} + g_4h_{n-2} + .. + g_1h_1 + g_2h_0)x^2 + ... + (g_0h_{n-1} + g_1h_{n-2} \\
& + ... + g_{n-1}h_0)x^{n-1} \pmod{x^n-1}.
\end{aligned}$$

Luego los coeficientes de esta última linea deben ser cero y como

$$\begin{aligned}
0 \equiv & (g_1g_2...g_{n-2}g_{n-1}g_0 \cdot h_{n-1}...h_1h_0) + (g_2g_3...g_{n-1}g_0g_1 \cdot h_{n-1}...h_1h_0)x \\
& + (g_3g_4...g_0g_1g_2 \cdot h_{n-1}...h_1h_0)x^2 + ... + (g_0g_1...g_{n-2}g_{n-1} \cdot h_{n-1}...h_1h_0)x^{n-1} \\
\equiv & (\sigma^{n-1}(g_0g_1...g_{n-2}g_{n-1}) \cdot h_{n-1}...h_1h_0) + (\sigma^{n-2}(g_0g_1...g_{n-2}g_{n-1}) \cdot h_{n-1}...h_1h_0)x \\
& + ... + ((\sigma^0(g_0g_1...g_{n-2}g_{n-1}) \cdot h_{n-1}...h_1h_0))x^{n-1} \pmod{x^n-1}.
\end{aligned}$$

se tiene que $h_{n-1}...h_1h_0$ es ortogonal a cada fila de la matriz generatriz de C dada por el Teorema (4.2.2), y así $h_{n-1}...h_1h_0$ es una palabra código de $C^\perp$.
Sabemos que $C^\perp$ es un código cíclico y como $h_{n-1}...h_1h_0$ es una de sus palabras código, también lo es $\sigma^{k+1}(h_{n-1}...h_1h_0)$ palabra a la que le corresponde por T $h_R(x)$. Así como $\mathrm{gr}(h_R(x)) = k$ y $h_R(x)|x^n-1$, el código cíclico de longitud n sobre F_q que tiene como polinomio generador a $h_r(x)$ tiene dimensión $n-k$ y además se tiene que el conjunto $\{h_R(x), xh_R(x), ..., x^{n-k-1}h_R(x)\}$ es linealmente independiente, y por lo tanto base de $T(C^\perp)$. Finalmente $h_R(x)$ genera $C^\perp$ y como $h_R(x)$ puede no ser mónico, el polinomio generador del código cíclico $C^\perp$ es $h_0^{-1}h_R(x)$. □

Corolario 4.2.11. *Sea C un $[n,k]$ código cíclico sobre F_q con polinomio generador $g(x)$. Sea $h(x) = h_0 + h_1x + ... + h_kx^k$ talque $h(x) = (x^n-1)/g(x)$. Entonces*

$$H = \begin{pmatrix}
h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & 0 & 0 & \cdots & 0 \\
0 & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & 0 & \cdots & 0 \\
0 & 0 & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & & \ddots & 0 \\
0 & 0 & \cdots & 0 & h_k & h_{k-1} & h_{k-2} & \cdots & h_0
\end{pmatrix}$$

es una matriz de control de paridad de C

Demostración. Por teorema anterior $h_R(x)$ genera $C^{\perp}$, tenemos por procedimientos análogos a la demostración del Teorema (4.2.2) que la matriz H definida arriba es una matriz generadora de $C^{\perp}$ y por lo tanto una matriz de control de paridad de C. □

Ejemplo 4.2.12. Considerando los códigos cíclicos y sus correspondientes polinomios generadores del Ejemplo (4.2.6)

(a) El polinomio generador dado fue x^2+2. Luego $h(x) = x^2+1$ y

$$h_0^{-1}h_R(x) = x^2+1 = h(x)$$

Así

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

y como vimos anteriormente $GH^{\top} = 0$.

(b) Si el código cíclico es generado por $x^4+x^3+x^2+1$, se tiene entonces que $h(x) = x^3+x^2+1$. Luego

$$h_0^{-1}h_R(x) = x^3+x+1$$

y así

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

es una matriz de paridad de C y se verifica que

$$GH^{\top} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

4.3. Codificación y decodificación de un código cíclico

El Teorema (4.2.2) nos facilita una matriz generadora de un código cíclico, la cual podemos llevarla a forma estándard mediante las operaciones de filas y columnas mencionadas en el Capítulo 2 y, al ser un código cíclico un código lineal, podemos

codificar y decodificar usando los algoritmos anteriormente presentados. Nuestro objetivo es mostrar algoritmos para codificar y decodificar mensajes pero utilizando elementos del ideal correspondiente al código cíclico en cuestión, i.e. utilizando polinomios, que es lo que caracteriza a este tipo de código.

Sea C un $[n,k]$, código cíclico sobre F_q, con polinomio generador $g(x)$. Resultará de utilidad usar una matriz generadora G de la forma

$$G = \left(\begin{array}{c|c} A & I_k \end{array} \right)$$

donde I_k es la matriz identidad de tamaño $k \times k$. Sabemos que una forma de conseguir esta matriz sería realizando operaciones entre filas y entre columnas, sin embargo, hay otra manera y es utilizando el polinomio generador de C, quizás tediosa para fines prácticos, pero útil a fines teóricos [12].
Lo que se hace es dividir a x^{n-k+i} por $g(x)$, con $i = 0, 1, ..., k-1$. De esta manera resulta

$$\begin{aligned} x^{x-k+i} &= q_i(x)g(x) + r_i(x) \\ x^{x-k+i} - r_i(x) &= q_i(x)g(x) \end{aligned}$$

donde $q_i(x), r_i(x) \in F_q[x]/x^n - 1$ y $\text{gr}(r_i(x)) < \text{gr}(g(x)) = n-k$ para $i = 0, 1, ..., k-1$. Podemos observar que los elementos que se corresponden con cada $x^{x-k+i} - r_i(x)$ por T son palabras código de C y que el conjunto $\{x^{x-k+i} - r_i(x)\}$ es linealmente independiente, lo que podemos verificar haciendo lo siguiente.
Sea $r_i(x) = \sum_{j=0}^{n-k-1} r_{i,j}x^j$, entonces formando la matriz G con los coeficientes de $x^{x-k+i} - r_i(x)$, resulta

$$G = \begin{pmatrix} -r_{0,0} & -r_{0,1} & \cdots & -r_{0,n-k-1} & 1 & 0 & \cdots & 0 \\ -r_{1,0} & -r_{1,1} & \cdots & -r_{1,n-k-1} & 0 & 1 & & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \ddots & 0 \\ -r_{k-1,0} & -r_{k-1,1} & \cdots & -r_{k-1,n-k-1} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

con filas linealmente independientes y así una matriz generadora de C.

Teorema 4.3.1. *Sea C un $[n,k]$ código cíclico sobre F_q, con $g(x)$ su polinomio generador. Entonces la palabra código c proveniente del mensaje m se corresponde con el polinomio $c(x) = q(x)g(x)$ donde $q(x)$ es el cociente de dividir $x^{n-k}m(x)$ por $g(x)$.*

Demostración. Sea $m(x) = m_0 + m_1x + ... + m_{k-1}x^{k-1}$. Hagamos $x^i x^{n-k} = q_i(x)g(x) + r_i(x)$, con $i = 0, 1, ..., k-1$. Multiplicando m_i miembro a miembro como sigue

$$\begin{aligned} m_0x^{n-k} &= m_0q_0(x)g(x) + m_0r_0(x) \\ m_1xx^{n-k} &= m_1q_1(x)g(x) + m_1r_1(x) \\ &\vdots \\ m_{k-1}x^{k-1}x^{n-k} &= m_{k-1}q_{k-1}(x)g(x) + m_{k-1}r_{k-1}(x) \end{aligned}$$

y sumando resulta

$$\begin{aligned} m(x)x^{n-k} =& g(x)\left[m_0q_0(x) + m_1q_1(x) + ... + m_{k-1}q_{k-1}(x)\right] \\ &+ \left[m_0r_0(x) + m_1r_1(x) + ... + m_{k-1}r_{k-1}(x)\right]. \end{aligned}$$

considerando la matriz generadora G como está arriba se puede verificar que

$$T(mG) = -m_0r_0(x) - m_1r_1(x) - ... - m_{k-1}r_{k-1}(x) + m(x)x^{n-k}$$

Desarrollando $-m_0r_0(x) - m_1r_1(x) - ... - m_{k-1}r_{k-1}(x) + m(x)x^{n-k}$ pero usando los elementos correspondientes en F_q^n se tiene

$$\begin{aligned} &m_0(-r_{0,0} - r_{0,1}... - r_{0,n-k-1}10..,0) + m_1(-r_{1,0} - r_{1,1}... - r_{1,n-k-1}010..,0) + ... \\ &+ m_{k-1}(-r_{k-1,0}... - r_{k-1,n-k-1}0..,01) \\ &= (-m_0r_{0,0} - m_1r_{1,0} - ... - m_{k-1}r_{k-1,0} \quad - m_0r_{0,1} - m_1r_{1,1} - ... - m_{k-1}r_{k-1,1} \quad ... \\ &- m_0r_{0,n-k-1} - m_1r_{1,n-k-1} - ... - m_{k-1}r_{k-1,n-k-1} \quad m_0 \ ... \ m_{k-1}) \end{aligned}$$

y esto último es mG. Así si llamamos $c(x) = -m_0r_0(x) - m_1r_1(x) - ... - m_{k-1}r_{k-1}(x) + m(x)x^{n-k}$, tenemos que la palabra código c proveniente del mensaje m se corresponde con el polinomio $c(x) = q(x)g(x)$ donde $q(x)$ es el cociente de dividir $x^{n-k}m(x)$ por $g(x)$. □

Ejemplo 4.3.2. Sea C el código cíclico de longitud 7, generado por $1 + x^2 + x^3$. Tenemos que $n = 7$, $k = 4$, entonces para encontrar nuestra matriz generadora deseada dividamos x^3, x^4, x^5 y x^6 por $1 + x^2 + x^3$.

$$\begin{aligned} x^3 &= (x^3 + x^2 + 1) + (x^2 + 1) \\ x^4 &= (x + 1)(x^3 + x^2 + 1) + (x^2 + x + 1) \\ x^5 &= (x^2 + x + 1)(x^3 + x^2 + 1) + (x + 1) \\ x^6 &= (x^3 + x^2 + x)(x^3 + x^2 + 1) + (x^2 + x) \end{aligned}$$

por lo tanto

$$G = \left(\begin{array}{ccc|cccc} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right).$$

Si por ejemplo codificamos el mensaje $m = 1001$ haciendo mG, se tiene que éste queda codificado como la palabra código $c = 1101001$.
Ahora $m(x) = 1 + x^3$ y $x^{n-k} = x^3$, por lo tanto

$$x^3m(x) = (x^3 + x^2 + x + 1)(x^3 + x^2 + 1) + (x + 1)$$

y así $c(x) = x^3m(x) - (x + 1) = (x^3 + x^2 + x + 1)(x^3 + x^2 + 1) = 1 + x + x^3 + x^6$

Otro ejemplo es codificar el polinomio $p(x) = x^3 + x^2$.

$$x^3 p(x) = (x^3 + 1)(x^3 + x^2 + 1) + (x^2 + 1)$$

por lo tanto $p(x)$ se codifica como $1 + x^2 + x^5 + x^6$. Queda como ejercicio para el lector que mostrar que el mensaje de F_q^k que se corresponde a $p(x)$ se codifica como la palabra código de C que se corresponde con $1 + x^2 + x^5 + x^6$.

Teniendo en cuenta la matriz $G = (A|I_k)$ dada arriba, por Teorema (2.3.13), una matriz de control de paridad para C es

$$H = (I_{n-k}|R)$$

donde I_{n-k} es la matriz identidad de tamaño $n - k$ y $R = -A^\top$.

Teorema 4.3.3. *Sea $H = (I_{n-k}|R)$ una matriz de control de paridad C un $[n, k]$ código cíclico sobre F_q, con $g(x)$ su polinomio generador. Sea $a(x)$ el polinomio de grado a lo sumo $n - 1$ correspondiente al elemento a de F_q^n, y $s(x)$ el polinomio correspondiente al síndrome s de a. Entonces $s(x)$ es $(a(x) mod\ (g(x)))$, i.e. $s(x)$ es el resto que resulta al dividir $a(x)$ por $g(x)$.*

Demostración. Recordemos que $x^{n-k+i} = q_i(x)g(x) + r_i(x)$ para $i = 0, 1, ..., k - 1$ y notemos que las i-ésima columna de R está formada por los coeficientes de $r_i(x)$. Sea $a(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1}$. Sabemos que

$$s = aH^\top = a_0 \cdot H_0 + a_1 \cdot H_1 + ... + a_{n-1} \cdot H_{n-1}$$

Con H_i la i-ésima columna de H.
Reemplazando por los elementos correspondientes de $F_q[x]/x^n - 1$ se tiene que

$$\begin{aligned} s(x) &= a_0 1 + a_1 x + .. + a_{n-k-1}x^{n-k-1} + a_{n-k}r_0(x) + a_{n-k+1}r_1(x) + ... + a_{n-1}r_{k-1}(x) \\ &= a_0 + a_1 x + .. + a_{n-k-1}x^{n-k-1} + a_{n-k}(x^{n-k} - g(x)q_0(x)) \\ &\quad + a_{n-k+1}(x^{n-k+1} - g(x)q_1(x)) + ... + a_{n-1}(x^{n-1} - g(x)q_{k-1}(x)) \\ &= a(x) - g(x)\left[a_{n-k}q_0(x) + a_{n-k+1}q_1(x) + ... + a_{n-1}q_{k-1}(x)\right]. \end{aligned}$$

Como $\mathrm{gr}(r_i(x)) \leq n - k - 1$ para $i = 0, 1, ..., k - 1$, se observa que

$$\mathrm{gr}(s(x)) \leq n - k - 1 < n - k = \mathrm{gr}(g(x))$$

por lo tanto $s(x)$ es el resto que resulta al dividir $a(x)$ por $g(x)$, i.e.
$s(x) \equiv a(x) \pmod{g(x)}$. □

En el Capítulo 2 definimos el síndrome de un elemento de $y \in F_q^n$ como $S(y) = yH^\top$. Consideremos $a(x)$ y $s(x)$ como en el teorema anterior. Con un abuso en la notación diremos que $S(a(x)) = s(x)$ y también que $wt(a(x)) = wt(a)$. Esto lo hacemos a fin de agilizar la lectura de los siguientes teoremas.

Teorema 4.3.4. *Sea C un $[n,k]$ código cíclico sobre F_q, con polinomio generador $g(x)$. Sea $s(x) = \sum_{i=0}^{n-k-1} s_i x^i$ el síndrome de $w(x) \in F_q[x]/x^n - 1$. Entonces el síndrome de $xw(x)$ es $xs(x) - s_{n-k-1}g(x)$.*

Demostración. Sabemos que $w(x) = q(x)g(x) + s(x)$, con $\mathrm{gr}(s(x)) \leq n-k-1$ o $s(x) = 0$. Para $s(x) = 0$ se cumple la tésis de forma trivial.
Veamos el caso en que $\mathrm{gr}(s(x)) \leq n-k-1$. Tenemos que

$$xw(x) = xq(x)g(x) + xs(x)$$

Si $\mathrm{gr}(s(x)) < n-k-1$, entonces $xs(x)$ es el resto de de $xw(x)$ ya que $\mathrm{gr}(xs(x)) < n-k$, y como, para este caso, $s_{n-k-1} = 0$ tiene que $S(w(x)) = xs(x) - 0g(x)$.
Para $\mathrm{gr}(s(x)) = n-k-1$, dividamos $xs(x)$ por $g(x)$ obteniendo el cociente $q_1(x)$ y el resto $t(x)$. Luego

$$xw(x) = q(x)g(x) + q_1(x)g(x) + t(x)$$

y como $\mathrm{gr}(xs(x)) = n-k$ y $g(x)$ mónico, se tiene que $q_1(x) = s_{n-k-1}$ y así $t(x) = xs(x) - s_{n-k-1}g(x)$. Entonces

$$xw(x) = (q(x) + s_{n-k-1})g(x) + (xs(x) - s_{n-k-1}g(x))$$

con $\mathrm{gr}(t(x)) < n-k$. De esto último se sigue la tésis. □

Teorema 4.3.5. *Sea C un código cíclico con polinomio generador $g(x)$. Supongamos que en una transmisión se recibe $w(x)$, con síndrome $s(x)$.*
Si $wt(s(x)) \leq \lfloor (d(C)-1)/2 \rfloor$, entonces $s(x)$ es el patrón de error de $w(x)$, i.e. $w(x)$ se decodifica como $w(x) - s(x)$.

Demostración. Sea s la palabra que se corresponde con $s(x)$ por T. Por hipótesis $wt(s) \leq \lfloor (d(C)-1)/2 \rfloor$, y como $S(w(x)) = s(x) = S(s(x))$, tenemos que s es el representante de la coclase en que se encuentra la palabra w que se corresponde con $w(x)$ por T. Por el algoritmo de la decodificación por síndrome dado en el Capítulo 2, tenemos que w se decodifica como $w - s$, lo que en polinomios es $w(x) - s(x)$. □

Observación 4.3.6. Aunque no se encuentre mencionado explícitamente en los teoremas, debemos recordar que las matrices generadoras y de control de paridad de C, un código cíclico, son los expuestos arriba, y tienen la forma $\left(G = (-A^\top | I_k)\right)$ y $(H = (I_{n-k}|A)$ respectivamente.

Decodificaremos mediante la decocificación por síndrome. El teorema anterior asegura que cuando $wt(s(x)) \leq \lfloor (d(C)-1)/2 \rfloor$, con $S(w(x)) = s(x)$, entonces $w(x)$ se decodifica como $w(x) - s(x)$. El problema está en que no siempre para un $w(x)$ recibido se cumple que su síndrome tenga peso a lo sumo $\lfloor (d(C)-1)/2 \rfloor$. Nuestro trabajo ahora es encontrar el patrón de error que llamaremos $e(x)$, no usando la

tabla presentada en el Capítulo 2, sino utilizando los teoremas anteriores y las siguientes consideraciones.

Sea C un $[n, k, d]$ código cíclico sobre F_q con polinomio generador $g(x)$. Consideremos $(H = (I_{n-k}|A)$ una matriz de control de paridad de C. Supongamos que se recibe una palabra $w(x)$ que tiene un patrón de error $e(x)$, donde $wt(e(x)) \leq \lfloor (d(C) - 1)/2 \rfloor$ y suponiendo que si e se corresponde a $e(x)$ pot T, entonces e tiene una cadena de $0s$, contados de forma cíclica, de longitud al menos k (e.g. 100120001 tiene una cadena de $0s$ de longitud 3, 00010100 tiene una cadena de $0s$ de longitud 5).
Sea $s(x) = \sum_{i=0}^{n-k-1} s_i x^i$ el síndrome de $w(x)$. Sabemos que $s(x) = S(s(x))$. Denotando con $\overline{w(x)}$ a $w(x)$ (mod $g(x)$) se tiene que $S(w(x)) = \overline{w(x)} = s(x) = \overline{S(s(x))}$ [14]. Luego

$$S(xw(x)) = \overline{xs(x) - s_{n-k-1}g(x)} = \overline{xs(x)}$$

Así

$$\begin{aligned} S(x^2w(x)) &= \overline{xS(xw(x))} = \overline{x}\overline{xs(x)} \\ &= \overline{x^2s(x)} \\ S(x^3w(x)) &= \overline{xS(x^2w(x))} = \overline{x}\overline{x^2s(x)} \\ &= \overline{x^3s(x)} \end{aligned}$$

y más general

$$S(x^iw(x)) = \overline{x^is(x)} \text{ para todo } i = 0, ..., n-1$$

Por otro lado, sea m el entero tal que $\sigma^m(e)$ tenga los valores que no son parte de la cadena de $0s$ en las primeras $n - k$ coordenadas. Por lo expuesto con anterioridad $S(x^me(x)) = \overline{x^ms(x)}$. Definimos $s^i(x)$ como el síndrome de $x^iw(x)$ con $i = 0, ..., n-1$. De esta manera $\overline{x^ms(x)} = s^m(x)$ y como $H = (I_{n-k}|A)$

$$wt(\sigma^m(e)) = wt(e) = wt(s^m(x)) \leq \lfloor (d(C) - 1)/2 \rfloor$$

Así existe un entero m tal que el síndrome de $x^mw(x)$ es menor o igual a $\lfloor (d(C) - 1)/2 \rfloor$.

Ahora, sea $s^m(x)$ el síndrome de $x^mw(x)$ tal que $s^m(x) \leq \lfloor (d(C) - 1)/2 \rfloor$, para algún $m = 0, ..., n - 1$. Sea $l_m = (s^m|0) \in F_q^n$ con s^m la palabra en F_q^{n-k} que se corresponde con $s^m(x)$. Teniendo en cuenta la matriz H que estamos usando, es claro que $S((s^m|0)) = s^m$. También se cumple que $\sigma^n((s^m|0)) = (s^m|0)$, que en polinomios resulta $x^nl_m(x) = l_m(x)$. Luego

$$S(x^nl_m(x)) = s^m(x) = \overline{x^ms(x)}$$

Por otro lado se tiene que $S(x^n l_m(x)) = \overline{x^n s^m(x)}$. Así

$$\begin{aligned}
\overline{x^m s(x)} &= \overline{x^n s^m(x)} \\
\overline{x^m}\,\overline{s(x)} &= \overline{x^n}\,\overline{s^m(x)} \\
\overline{s(x)} &= \overline{x^{n-m}}\,\overline{s^m(x)} \\
s(x) &= \overline{x^{n-m} s^m(x)}
\end{aligned}$$

Notemos que $x^{n-m}s^m(x) \mod (x^n - 1)$ se corresponde a $\sigma^{n-m}(s^m)$ por T. También se cumple que $\sigma^{n-m}(s^m) = \sigma^{n-m}(l_m)$ y por lo tanto $S(\sigma^{n-m}(l_m)) = s$, lo que en polinomios es

$$S(x^{n-m} l_m(x)) = s(x).$$

Como $wt(l_m(x)) = wt(x^{n-m} l_m(x)) = wt(e(x)) \leq \lfloor (d(C) - 1)/2 \rfloor$ y al ser $S(x^{n-m} l_m(x)) = S(e(x)) = s(x)$ concluimos que $\sigma^{n-m}(l_m)$ y e están en la misma coclase al tener el mismo peso y poder ser representantes de su coclase, concluimos que $x^{n-m} l_m(x) = e(x)$, o lo que es lo mismo, $e(x)$ es $x^{n-m}s^m(x) \mod (x^n - 1)$.

El algoritmo de decodificación es el siguiente:

Sea C un $[n, k, d]$ código cíclico sobre F_q, con polinomio generador $g(x)$. Sea $w(x)$ la palabra que se recibe luego de una transmisión y $e(x)$ el patrón de error donde $wt(e(x)) \leq \lfloor (d(C) - 1)/2 \rfloor$ y $e(x)$ tiene una cadena de $0s$ de longitud al menos k contados cíclicamente.

Paso 1 Calcular los síndromes $s^i(x)$ de $x^i w(x)$ con $i = 0, 1, ..., n - 1$.

Paso 2 Encontrar m tal que el síndrome $s^m(x)$ de $x^m w(x)$ tenga peso menor o igual que $\lfloor (d(C) - 1)/2 \rfloor$.

Paso 3 Calcular $e(x)$ como el resto de dividir $x^{n-m}s^m(x)$ por $x^n - 1$. Decodificar $w(x)$ como $w(x) - e(x)$.

El ejemplo que sigue contiene cálculos que relativos a encontrar el resto de una división de polinomios, que es lo que debemos hacer tanto para codificar mensajes como para decodificarlos. El lector puede ayudarse con el software GeoGebra (con el comando Resto) o con el divisor de polinomios que se encuentra en la página web `https://es.planetcalc.com/7717/`, muy útil y el que se recomienda para verificar los cálculos.

Ejemplo 4.3.7. Sea C un código cíclico binario de longitud 15 generado por $g(x) = 1 + x^4 + x^6 + x^7 + x^8$.
Determinemos una matriz generadora y una de control de paridad de C.

Sea $r(f(x), g(x))$ el resto que resulta de dividir $f(x)$ por $g(x)$, entonces

$$r(x^8, g(x)) = 1 + x^4 + x^6 + x^7 \qquad r(x^{12}, g(x)) = x + x^3 + x^4 + x^5$$
$$r(x^9, g(x)) = 1 + x + x^4 + x^5 + x^6 \qquad r(x^{13}, g(x)) = x^2 + x^4 + x^5 + x^6$$
$$r(x^{10}, g(x)) = x + x^2 + x^5 + x^6 + x^7 \qquad r(x^{14}, g(x)) = x^3 + x^5 + x^6 + x^7$$
$$r(x^{11}, g(x)) = 1 + x^2 + x^3 + x^4$$

Luego

$$G = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & I_7 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & \end{array}\right)$$

es una matriz generadora y

$$H = \left(\begin{array}{c|ccccccc} & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ I_8 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

una matriz de control de paridad de C. Notemos que H tiene 5 columnas linealmente dependientes, y cualquier conjunto de 4 columnas son linealmente independientes. Luego por Teorema (2.3.15) $d(C) = 5$ y así puede corregir a lo sumo 2 errores. Supongamos que luego de una transmisión

(a) recibimos a $w_1(x) = x^4 + x^8 + x^{10} + x^{11} + x^{12} + x^{14}$. Calculamos los $s^i(x)$.

$$s^0(x) = x^3 + x^5 + x^6 + x^7 \qquad s^1(x) = 1$$

$x^{15-1}s^1(x) = x^{14}$. $e(x) = x^{14}$ ya que x^{14} es él mismo módulo $x^{15}-1$. Finalmente decodificamos $w_1(x)$ como

$$w_1(x) - e(x) = x^4 + x^8 + x^{10} + x^{11} + x^{12}.$$

(b) recibimos $w_2(x) = 1 + x^3 + x^5 + x^{12} + x^{14}$.

$$s^0(x) = 1 + x + x^3 + x^4 + x^5 + x^6 + x^7$$
$$xw_2(x) = 1 + x + x^4 + x^6 + x^{13} \qquad s^1(x) = 1 + x + x^2 + x^5$$
$$x^2w_2(x) = x + x^2 + x^5 + x^7 + x^{14} \qquad s^2(x) = x + x^2 + x^3 + x^6$$
$$x^3w_2(x) = 1 + x^2 + x^3 + x^6 + x^8 \qquad s^3(x) = x^2 + x^3 + x^4 + x^7$$
$$x^4w_2(x) = x + x^3 + x^4 + x^7 + x^9 \qquad s^4(x) = 1 + x^3 + x^5 + x^6 + x^7$$
$$x^5w_2(x) = x^2 + x^4 + x^5 + x^8 + x^{10} \qquad s^5(x) = 1 + x$$

Así $m = 5$. $x^{10}s^5(x) = x^{10}(1+x) = x^{10} + x^{11} = e(x)$. Luego decodificamos $w_2(x)$ como

$$w_2(x) - e(x) = 1 + x^3 + x^5 + x^{10} + x^{11} + x^{12} + x^{14}$$

(c) recibimos $w_3(x) = x^3 + x^8 + x^{10} + x^{12} + x^{13}$.

$$\begin{array}{ll} & s^0(x) = 1 + x^4 + x^5 + x^6 \\ xw_3(x) = x^4 + x^9 + x^{11} + x^{13} + x^{14} & s^1(x) = x + x^5 + x^6 + x^7 \\ x^2w_3(x) = 1 + x^5 + x^{10} + x^{12} + x^{14} & s^2(x) = 1 + x^2 + x^4 \\ x^3w_3(x) = 1 + x + x^6 + x^{11} + x^{13} & s^3(x) = x + x^3 + x^5 \\ x^4w_3(x) = x + x^2 + x^7 + x^{12} + x^{14} & s^4(x) = x^2 + x^4 + x^6 \\ x^5w_3(x) = 1 + x^2 + x^3 + x^8 + x^{13} & s^5(x) = x^3 + x^5 + x^7 \\ x^6w_3(x) = x + x^3 + x^4 + x^9 + x^{14} & s^6(x) = 1 + x^7 \end{array}$$

Luego $x^9s^6(x)x^9(1+x^7) = x^9 + x^{16}$ que módulo $x^{15} - 1$ es $x + x^9 = e(x)$. Decodificamos $w_3(x)$ como

$$w_3(x) - e(x) = x + x^3 + x^8 + x^9 + x^{10} + x^{12} + x^{13}.$$

Observación 4.3.8. $s^i(x)$ en el ejemplo anterior se calculó como $r(x^iw(x), g(x))$. Sin embargo anteriormente concluimos que $S(x^iw(x)) = \overline{x^is(x)}$ para todo $i = 0, ..., n-1$, así podríamos haber calculado $s^i(x)$ a partir de $s^0(x)$ calculando $\overline{x^is(x)}$ para todo $i = 0, ..., n-1$.

Teorema 4.3.9. *Sean C_1 y C_2 dos códigos cíclicos de F_q^n con polinomios generadores $g_1(x)$ y $g_2(x)$ respectivamente. Entonces, $C_1 \subseteq C_2$ si y sólo si $g_2(x)|g_1(x)$.*

La demostración de este teorema se deja como ejercicio para el lector.

Hay muchos códigos cíclicos famosos como el cídigo BCH, el Goppa, Reed-Solomon, residuos cuadráticos. En el capítulo anterior estudiamos los códigos Hamming y Golay. Los códigos G_{23} y G_{11} son códigos cíclicos. Consideremos la siguiente factorización de $x^{23} - 1$ en F_2

$$x^{23} - 1 = (x-1)(x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1)(x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1)$$

G_{23} tiene como polinomio generador a $x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1$.
De la misma forma, una factorización de $x^{11} - 1$ en F_3 es

$$x^{11} - 1 = (x-1)(x^5 + x^4 - x^3 + x^2 - 1)(x^5 - x^3 + x^2 - x - 1)$$

y el polinomio generador de G_{11} es $x^5 + x^4 - x^3 + x^2 - 1$.

El Hamming q-ario no siempre lo es, sin embargo, se tiene lo siguiente.

Teorema 4.3.10. *El código $Ham(r,2)$ es equivalente a un código cíclico.*

Ejemplo 4.3.11. (a) Para $r = 2$ una matriz de control de paridad es

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

Esta matriz es equivalente a la matriz

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

que es generadora de un código cíclico. Observando las columnas de esta última matriz podemos afirmar que es una matriz de control de paridad de $Ham(2,2)$ es decir del código de repetición $[3,1]$. Para determinar su polinomio generador identificamos a $h_R(x) = 1 + x = h(x)$. Como
$x^3 - 1 = (x+1)(x^2+x+1)$ concluimos que $g(x) = x^2 + x + 1$ es el polinomio generador de $Ham(2,2)$.

(b) Para $r = 3$ una matriz de control de paridad es

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

que es equivalente a la matriz

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

El código $Ham(3,2)$ es entonces el código cíclico $[7,4,3]$.
$h_R(x) = 1 + x^2 + x^3 + x^4$ por lo tanto $h(x) = 1 + x + x^2 + x^4$, así resulta que $x^7 - 1 = (1+x+x^2+x^4)(x^3+x+1)$. Finalmente concluimos que $g(x) = x^3+x+1$ es el polinomio generador de $Ham(3,2)$.

(c) Para $r = 4$ una matriz de control de paridad es

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

que es equivalente a la matriz

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

Por lo tanto $Ham(4,2)$ es el código cíclico $[15,11,3]$.
$h_R(x) = 1 + x + x^2 + x^3 + x^5 + x^7 + x^8 + x^{11}$ y su recíproco es
$h(x) = 1 + x^3 + x^4 + x^6 + x^8 + x^9 + x^{10} + x^{11}$ y como
$x^{15} - 1 = (1 + x^3 + x^4 + x^6 + x^8 + x^9 + x^{10} + x^{11})(x^4 + x^3 + 1)$ concluimos que
$g(x) = x^4 + x^3 + 1$ es el polinomio generador de $Ham(4,2)$.

4.4. Comentarios finales

Sea C el código cíclico binario de longitud 15 generado por $g(x) = 1+x^4+x^6+x^7+x^8$ del Elemplo (4.3.7). Por lo visto en el ejemplo mencionado, C es un código $[15,7,5]$ con matriz generadora y matriz de control de paridad G y H respectivamente, donde

$$G = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & I_7 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & \end{array}\right)$$

y

$$H = \left(\begin{array}{c|ccccccc} & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ I_8 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Utilizando lo aprendido en el Caítulo 2, C tiene $2^7 = 128$ elementos, y $|F_2^{15}| = 2^{15} = 32768$. Luego existen $32768 : 128 = 256$ coclases distintas. Para realizar la decodificación por síndome necesitamos los representantes que cumplen con tener peso menor o igual a $\lfloor \frac{5-1}{2} \rfloor = 2$. Para determinar las coclases debemos tener primero todos las palabras códigos y para hacerlo debemos codificar cada uno de los 128 mensajes posibles de longitud 7, lo que implica multiplicar cada una de estos 128 mensajes por G. Esto ya es un trabajo arduo y aún falta determinar las coclases y seleccionar de cada una un representante.
Por otro lado, es posible construir la tabla de Representante-Síndrome calculando el síndrome de aquellas palabras que tengan peso menor o igual a 2. La cantidad de palabras que cumplen con esta condicion son $\binom{15}{1}\binom{15}{2} = 120$. Así debemos realizar a lo sumo 120 multiplicaciones, una por cada palabra, por H lo cuál no

resulta eficiente.
Entonces, trabajar el código cíclico C con elementos del anillo $F_2[x]/x^{15}-1$ trae algunas ventajas, por ejemplo

- para codificar un mensaje, no necesitamos la matriz generadora del código para determinar la palabra código que le corresponde al mensaje, sólo necesitamos el polinomio generador de C.

Es cierto, que la decodificación dada para C en $F_2[x]/x^{15}-1$ es en esencia la decodificación por síndrome dada en el Capítulo 2, sin embargo, observemos que en $F_2[x]/x^{15}-1$ contamos con algunas ventajas:

- el síndrome de un polinomio $a(x) \in F_2[x]/x^{15}-1$ es el resto de dividir $a(x)$ por $g(x)$. Así, para calcular el síndrome de un polinomio, no necesitamos la matriz de control de paridad, sólo necesitamos el polinomio generador de C,
- para determinar el error $e(x)$ producido en alguna transmisión, debemos realizar a lo sumo 15 multiplicaciones entre polinomios y para cada polinomio que resultan de estas multiplicaciones determinar su síndrome.

Con esto concluimos que realizar la decodificación por síndrome de un código cíclico q-ario $[n,k,d]$ resulta más eficiente en $F_q[x]/x^n-1$ que en F_q^n.

Conclusiones

Para el desarrollo de esta Introducción a la Teoría de Códigos utilizamos saberes del álgebra lineal, de estructuras algebraicas y de probabilidad. Esta Teoría es un ejemplo claro de la aplicación de la matemática a problemas prácticos de la vida real.

Para un estudiante es importante conocer estas aplicaciones prácticas para complementar lo aprendido en su curricula. Fue gratificante descubrir en este Teoría, una aplicación de conceptos tan abstractos como grupos, anillos; de elementos como los polinomios cuya utilidad es tan cuestionada por los egresados de la educación media.

Hemos presentado a la familia de los códigos Hamming y Golay, con sus características y su decodificación particular en cada caso. De acuerdo a lo leído podemos afirmar que el gran desafío de la Teoría de Códigos es la búsqueda de códigos optimos, de buenos códigos, pero también el desafío consiste en crear nuevas técnicas de decodificación.

Para el lector que desea continuar con el estudio de esta introducción, se recomienda la lectura de [4] para completar el estudio de códigos cíclicos y posteriormente estudiar la familia de los códigos Reed-Solomon construidos a partir de códigos cíclicos. Estos códigos fueron muy usados en la últimas décadas para corregir errores ocurridos en los dispositivos de almacenamiento, como en los discos compactos; en las comunicaciones inalámbricas; comunicaciones satelitales y televisión digital. Por su importancia, considero que es conveniente profundizar su estudio en un trabajo futuro.

A. Apéndice

Definición A.0.1. Un *anillo* es un conjunto no vacío R junto con dos operaciones binarias (usualmente escritas como suma y multiplicación) tales que

(a) $(R, +)$ es un grupo abeliano,

(b) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para todo $a, b, c \in R$,

(c) $a \cdot (b + c) = a \cdot b + a \cdot c$ y $(a + b) \cdot c = a \cdot c + b \cdot c$ para todo $a, b, c \in R$.

Un anillo R suele escribirse como la terna $(R, +, \cdot)$ donde $+$ y $\cdot$ son las operaciones que le dan la estructura de anillo siguiendo la notación usada en la definición anterior.

Definición A.0.2. Sea R un anillo

(a) R es un *anillo conmutativo* si satisface que $ab = ba$, pata todo $a, b \in R$.

(b) R es un *anillo con identidad* si contiene un elemento 1_R tal que $1_R a = a1_R$ para todo $a \in R$.

Ejemplo A.0.3. $\mathbb{Z}$ y $\mathbb{Z}_n$ son anillos de polinomios conmutativos con identidad bajo la suma y el producto usual y la suma y el producto módulo n respectivamente. El conjunto de los números enteros pares con la suma y multiplicación usual es un anillo conmutativo.

Definición A.0.4. Un elemento a del anillo R es *inversible* o es una *unidad* si existe un elemento $c \in R$ tal que $ca = 1_R = ac$.

Definición A.0.5. Un anillo con identidad R, con $1_R \neq 0$, es llamado *anillo de división* si cada elemento no nulo de R es una unidad. Un *cuerpo* es un anillo de división conmutativo.

Ejemplo A.0.6. $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son algunos ejemplos de cuerpos. También lo es $\mathbb{Q}(\sqrt{2})$ con

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}.$$

Definición A.0.7. Sea R un anillo. S es un *subanillo* de R si es un subconjunto no vacío de R que es a la vez un anillo respecto las operaciones de R.

Definición A.0.8. Sea R un anillo. Un *ideal* I de R es un subanillo de R que satisface que

(a) si $r \in R$ y $x \in I$, entonces $rx \in I$

(b) si $r \in R$ y $x \in I$, entonces $xr \in I$

Definición A.0.9. Un ideal I de un anillo R es un *ideal principal* si existe un elemento $g \in I$ tal que $I = \langle g \rangle = \{gr : \ r \in R\}$. Si todos los ideales de R son ideales principales, R recibe el nombre de *anillo ideal principal*.

Teorema A.0.10. *Un subconjunto no vacío I de R es un ideal si y sólo si para todo $a, b \in I$ y $r \in R$ se cumple que*

(a) $a - b \in I$ *y*

(b) $ra \in I$ *y* $ar \in I$.

Ejemplo A.0.11. (a) El conjunto I de todos los múltiplos de 3, es un ideal del anillo $\mathbb{Z}$. De forma general, sea R cualquier anillo conmutativo y $c \in R$, entonces el conjunto $I = \{r \cdot c | \, r \in R\}$ es un ideal de R, con $\cdot$ la segunda operación que le da la estructura de anillo a R.

(b) El conjunto $I = \{(k, 0) | \, k \in \mathbb{Z}\}$ es un ideal del anillo $\mathbb{Z} \times \mathbb{Z}$.

Definición A.0.12. Sea D un subanillo del anillo $(R, +, \cdot)$ y $a, b \in R$. a es congruente a derecha a b módulo D, denotado como $a \equiv_d b \pmod{D}$ si $a - b \in D$. a es congruente a izquierda a b módulo D, denotado como $a \equiv_i b \pmod{D}$ si $-a + b \in D$.

Teorema A.0.13. *Sea D un subanillo del anillo $(R, +, \cdot)$.*

(a) La congruencia (a izquierda y derecha) módulo D es una relación de equivalencia.

(b) La clase de equivalencia de $a \in R$ bajo la congruencia a derecha [a izquierda] módulo D es el conjunto $D + a = \{d + a : \ d \in D\}$ $[a + D = \{a + d : \ d \in D\}]$.

Demostración. (a) Se deja como ejercicio para el lector.

(b) La clase de equivalencia de $a \in R$ módulo D respecto de la congruencia a derecha es

$$\begin{aligned} \{x \in R : \ x \equiv a\} &= \{x \in R : \ x - a \in D\} \\ &= \{x \in R : \ x - a = d \in D\} \\ &= \{x \in R : \ x = d + a, \, d \in D\} \\ &= \{d + a : \ d \in D\}. \end{aligned}$$

De forma análoga se demuestra con la congruencia a izquierda.

□

Sea D un subanillo de R. Los conjuntos $D+a$ y $a+D$ se llaman coclases a derecha y a izquierda, respectivamente, de D en R. Como R es un anillo la suma es conmutativa y así se tiene que $D + a = a + D$. El conjunto R/D es el conjunto de todas las coclases de D en R y se llama conjunto cociente.
De forma general, sea A un conjunto, y sea ~una relacón binaria de equivalencia en A, entonces el conjunto de todas las clases de equivalencias (coclases) es llamado conjunto cociente.

Teorema A.0.14. *Sea R un anillo y D un ideal de R. Entonces R/D es un anillo bajo las operaciones*

$$\textit{Adición: } (a + D) + (b + D) = (a + b) + D$$

$$\textit{Multiplicación: } (a + D) \cdot (b + D) = (a \cdot b) + D.$$

Ejemplo A.0.15. Sea el ideal $I = \{3r | \, r \in \mathbb{Z}\}$ de $\mathbb{Z}$. El anillo cociente $\mathbb{Z}/I$ es

$$\mathbb{Z}/I = \{0 + I, 1 + I, 2 + I\}$$

Sea F un cuerpo. El conjunto de todos los polinomios sobre el cuerpo F, denotado como $F[x]$, donde

$$F[x] = \{a_0 + a_1x + ... + a_nx^n : \ a_i \in F, \ n \geq 0\},$$

tiene la estructura de anillo bajo la suma y multiplicación usual de polinomios.
Dado un polinomio no nulo $a(x)$, con

$$a(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n \quad \text{con } a_n \neq 0,$$

recordemos que n es el grado del polinomio $a(x)$ y se denota como $\text{gr}(r(x))$. Por convención, se dice que el grado del polinomio nulo es $-\infty$ [15], sin embargo, en este trabajo diremos que el polinomio nulo es aquel que no tiene grado.

Teorema A.0.16 (Algoritmo de la división de polinomios)**.** *Sea F un cuerpo y $f(x), g(x) \in F[x]$, con $g(x) \neq 0$. Entonces existe un único par de polinomios $q(x)$ y $r(x)$ tales que*

$$f(x) = q(x)g(x) + r(x)$$

donde $r(x) = 0$ ó $gr(r(x)) < gr(g(x))$.

Definición A.0.17. Sea F un cuerpo y $f(x), g(x) \in F[x]$, con $g(x) \neq 0$. Diremos que

(a) $g(x)$ divide a $f(x)$ (o que $g(x)$ es factor de $f(x)$) escrito como $g(x)|f(x)$, si $f(x) = h(x)g(x)$ para algún $h(x) \in F[x]$.

(b) $f(x)$ es *reducible* si $f(x) = h(x)q(x)$ donde $h(x), q(x) \in F[x]$ y el grado de $h(x)$ y $q(x)$ menores que el grado de $f(x)$. Si $f(x)$ no es reducible se llama *irreducible*.

(c) dos polinomios $q(x)$ y $h(x)$ de $F[x]$ son congruentes módulo $g(x)$, si $g(x)$ divide a $q(x) - h(x)$.

Teorema A.0.18. *Sea $g(x) \in F[x]$. La congruencia módulo $g(x)$ es una relación de equivalencia en $F[x]$.*

Observemos que, por el algoritmo de la división cualquier polinomio $f(x) \in F[x]$ es congruente módulo $g(x)$ a un único polinomio $r(x)$ con $\text{gr}(r(x)) < \text{gr}(g(x))$, donde $r(x)$ es el resto de dividir $f(x)$ por $g(x)$. Así $F[x]/g(x)$ es el conjunto de los polinomios de $F[x]$ con grado menor al grado de $g(x)$ unión el polinomio nulo. Estos polinomios son las clases de equivalencia, cada uno es el representante de su clase de equivalencia. Si F es finito entonces $F[x]/g(x)$ también es finito.

Sea $g(x) \in F[x]$, con F un cuerpo finito. Escribamos a $F[x]/g(x)$ como

$$F[x]/g(x) = \{r_1(x), r_2(x), ..., r_m(x)\},\ m \in \mathbb{N}$$

Definimos en $F[x]/g(x)$ las siguientes operaciones:

Adición: $r_j(x) \oplus r_i(x)$ es el resto de dividir $r_j(x) + r_i(x)$ por $g(x)$
Multiplicación: $r_j(x) \odot r_i(x)$ es el resto de dividir $r_j(x) \cdot r_i(x)$ por $g(x)$.

Bajo estas operaciones que acabamos de definir, $F[x]/g(x)$ es un anillo, llamado anillo de polinomios módulo $g(x)$.

Observación A.0.19.

1. Un conjunto cociente que tiene una estructura de anillo, se llama anillo cociente.

2. Algunas bibliografías definen anillo cociente a partir de un anillo R y la congruencia módulo I, con I un ideal de R.
   Consideremos $g(x)$ un elemento del anillo $F[x]$ y $I = \langle g(x) \rangle = \{a(x)g(x) : a(x) \in F[x]\}$ un ideal de $F[x]$. Sea $f(x), r(x) \in F[x]$ tal que $f(x)$ sea congruente con $r(x)$ módulo I. Esto significa que $f(x) - r(x) \in I$ es decir, existe un $a(x) \in F[x]$ talque
   $$f(x) - r(x) = a(x)g(x).$$
   Esta última expresión significa que $f(x) - r(x)$ es divisible por $g(x)$ que significa que $f(x)$ y $r(x)$ son congruentes módulo $g(x)$. Entonces la congruencia módulo I induce la congruencia módulo $g(x)$.

Así los anillos cocientes $F[x]/\langle x^n-1\rangle$ y $F[x]/x^n-1$ son iguales en cuanto a los elementos pertenecientes a cada uno de ellos, sin embargo distintos en cuanto a la relación de equivalencia que los definen.

Ejemplo A.0.20. (a) El anillo cociente $F_2[x]/x^3-1$ es

$$F_2[x]/x^3-1=\{0,1,1+x,x+x^2,1+x^2,x,x^2,1+x+x^2\}.$$

Cualquier otro polinomio de F_2x de grado mayor o igual a 3 pertenece a alguna de estas ocho clases de equivalencia. Por ejemplo, sea $x^5+x^4+x^2+1\in F_2[x]$. Tenemos que

$$x^5+x^4+x^2+1=(x^2+x)(x^3-1)+(1+x).$$

Entonces $x^5+x^4+x^2+1$ es congruente con $1+x$ módulo x^3-1, i.e. $x^5+x^4+x^2+1$ pertenece a la clase de equivalencia de $1+x$.

(b) El anillo cociente $F_2[x]/x^2+x+1$ es

$$F_2[x]/x^2+x+1=\{0,1,x,x+1\}$$

Teorema A.0.21. *Sea F un cuerpo y $g(x)\in F[x]$ un polinomio no constante. Si $f(x)\in F[x]$ y $g(x)$ son primos relativos, entonces $f(x)$ módulo $g(x)$ es una unidad en $F[x]/g(x)$.*

Demostración. Si $f(x)$ y $g(x)$ son primos relativos, entonces existen $s(x),t(x)\in F[x]$ tales que

$$1=s(x)f(x)+t(x)g(x),$$

luego $s(x)f(x)-1=-t(x)g(x)$.
De esto último se tiene que $s(x)f(x)-1\equiv 0 \mod g(x)$ y así $s(x)f(x)\equiv 1 \mod g(x)$. Esto quiere decir que $f(x)$ módulo $g(x)$ es una unidad en $F[x]/g(x)$ □

Teorema A.0.22. *Sea $g(x)$ un polinomio de grado mayor o igual a 1 sobre un cuerpo F. Entonces $F[x]/g(x)$ es un cuerpo si y sólo si $g(x)$ es irreducible en $F[x]$.*

Demostración. $\Rightarrow$) Supongamos por el absurdo que $g(x)$ es un polinomio reducible en $F[x]$. Así existen dos polinomios $a(x),b(x)\in F[x]$ con $0<\text{gr}(a(x))<\text{gr}(g(x))$ y $0<\text{gr}(b(x))<\text{gr}(g(x))$, tales que $g(x)=a(x)b(x)$. Como $a(x)$ y $b(x)$ tienen grados menores al grado de $g(x)$ entonces pertenecen a $F[x]/g(x)$. Por otro lado tenemos que $g(x)\equiv 0 \pmod{g(x)}$, por lo tanto se tiene que $0=a(x)b(x)$ y como $F[x]/g(x)$ es un cuerpo $a(x)=0$ o $b(x)=0$. Absurdo, pues $0<\text{gr}(a(x))<\text{gr}(g(x))$ y $0<\text{gr}(b(x))<\text{gr}(g(x))$, y no puede ocurrir que $a(x)=0$ ni $b(x)=0$ ya que el polinomio nulo no tiene grado. Como el absurdo se produjo al suponer que $g(x)$ es un polinomio reducible en $F[x]$, concluimos que $g(x)$ es irreducible en $F[x]$.

$\Leftarrow$) Demostrar que $F[x]/g(x)$ es un anillo (conmutativo con unidad) queda como ejercicio para el lector. Demostraremos a continuación que cada elemento no nulo de $F[x]/g(x)$ es una unidad.
Sea $f(x) \in F[x]/g(x)$, con $f(x) \not\equiv 0 \bmod g(x)$, y $m(x) = \text{mcd}(f(x), g(x))$. Así $m(x)|f(x)$ y $m(x)|g(x)$. Como $g(x)$ es irreducible, $m(x) = 1$ ó $m(x) = kg(x)$, con $k \in F$.
Si $m(x) = kg(x)$ entonces $kg(x)|f(x)$, i.e. $f(x) \equiv 0 \bmod g(x)$, lo que es falso. Por lo tanto $m(x) = 1$. Por teorema anterior, $f(x)$ es una unidad en $F[x]/g(x)$, y como $f(x)$ un elemento arbitrario, no nulo en $F[x]/g(x)$, concluimos que cada elemento no nulo de este anillo es una unidad y así $F[x]/g(x)$ es un cuerpo. □

Ejemplo A.0.23. $F_2[x]/x^2 + x + 1$ es un cuerpo. El lector puede verificar esto ayudandose de las siguientes tablas

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

(A.0.1)

Definición A.0.24. Sea F un cuerpo. La *característica* de F es el menor entero positivo p tal que $p \cdot 1 = 0$, donde 1 es la identidad de la multiplicación en F. Si este p no existe, diremos que la caracteristica de F es 0.

Ejemplo A.0.25. Los cuerpos $\mathbb{Z}_2$ y $\mathbb{Z}_3$ tienen características 2 y 3 respectivamente. de forma general, si p es un número primo, entonces el cuerpo $\mathbb{Z}_p$ es un cuerpo con característica p.
Los cuerpos $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ tienen caracteristica 0.

Teorema A.0.26. *La característica de un cuerpo es* 0 *ó un número primo.*

Demostración. Sea F un cuerpo. Tenemos que 1 no es la caracteristica del cuerpo ya que $1 \cdot 1 = 1 \neq 0$.
Supongamos que la característica de F es p y que p es un número compuesto. Sea $p = mn$ con n y m números enteros tales que $1 < n, m < p$. Llamamos $a = n \cdot 1$ y $b = m \cdot 1$, donde 1 es el la identidad de la multiplicación de F. Entonces

$$a \cdot b = (n \cdot 1)(m \cdot 1) = \left(\sum_{i=1}^{n} 1\right)\left(\sum_{j=1}^{m}\right) = (mn) \cdot 1 = p \cdot 1 = 0$$

Luego debe ocurrir que $a = 0$ o $b = 0$, i.e. $n \cdot 1 = 0$ o $m \cdot 1 = 0$. Absurdo, pues $1 < n, m < p$. Este absurdo provino de suponer que p es un número compuesto. Así concluimos que p es un número primo. □

Los siguientes teoremas nos dicen por un lado que si un cuerpo F es finito, entonces la cantidad de elementos que tiene es una potencia de un número primo, y por otro lado, que para cualquier número primo p y cualquier entero positivo n existe un cuerpo con p^n elementos. Estos teoremas no serán demostrados puesto que son demostraciones extensas y el último teorema requiere de elementos de la teoría de cuerpos finitos que no se considera oportuno estudiar en este trabajo.

Teorema A.0.27. *Un cuerpo finito F con caracteristica p contiene p^n elementos, para algun entero $n \geq 1$.*

Teorema A.0.28. *Para cualquier número primo p y cualquier entero $n \geq 1$, existe un único cuerpo finito de p^n elementos.*

Bibliografía

[1] Claude E Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[2] Richard E Blahut. *Algebraic codes for data transmission*. Cambridge university press, 2003.

[3] W Cary Huffman and Vera Pless. *Fundamentals of error-correcting codes*. Cambridge university press, 2010.

[4] San Ling and Chaoping Xing. *Coding theory: a first course*. Cambridge University Press, 2004.

[5] Raymond Hill. *A first course in coding theory*. Oxford University Press, 1986.

[6] Anton Betten, Michael Braun, Harald Fripertinger, Adalbert Kerber, Axel Kohnert, and Alfred Wassermann. *Error-correcting linear codes: Classification by isometry and applications*, volume 18. Springer Science & Business Media, 2006.

[7] Paul Garrett. *The Mathematics of coding theory*. Prentice-Hall, Inc., 2003.

[8] Wade Trappe. *Introduction to cryptography with coding theory*. Pearson Education India, 2006.

[9] Steven Roman. *Coding and information theory*, volume 134. Springer Science & Business Media, 1992.

[10] DG Hoffman, DA Leonard, CC Lindner, KT Phelps, CA Rodger, and JR Wall. Coding theory: The essentials, mercel dekker, 1992.

[11] Sanne van Alebeek. Comparing bounds on binary error-correcting codes. 2017.

[12] D J Baylis. *Error Correcting Codes: A Mathematical Introduction*. Routledge, 2018.

[13] Ricardo A Podestá. Introducción a la teoría de códigos autocorrectores. *Available: http://www2. famaf. unc. edu. ar/publicaciones/documents/serie_c/CMat35-3. pdf*, 2006.

[14] Francisco J Marín Ruiz. Técnicas algebraicas en la decodificación de códigos correctores de errores, 2018.

[15] Thomas W Hungerford. Graduate texts in mathematics: Algebra, 1974.