# CÓDIGOS DE CORRECCIÓN Y DETECCIÓN DE ERRORES

EDWIN ELIU MARTÍNEZ
MÓNICA JHOANA MESA MAZO

UNIVERSIDAD DEL CAUCA
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN
DEPARTAMENTO DE MATEMÁTICAS
POPAYÁN
2007

# CÓDIGOS DE CORRECIÓN Y DETECCIÓN DE ERRORES

EDWIN ELIU MARTÍNEZ
MÓNICA JHOANA MESA MAZO

**TRABAJO DE GRADO**
**En la modalidad trabajo de seminario presentado como requisito parcial para optar al título de matemático**

**Director**
**Dr. CARLOS ALBERTO TRUJILLO SOLARTE**

UNIVERSIDAD DEL CAUCA
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN
DEPARTAMENTO DE MATEMÁTICAS
POPAYÁN
2007

## Nota de aceptación

_____

_____

_____

Director

_____

**Doctor** Carlos Alberto Trujillo Solarte

**Comité evaluador**

_____

**Magister** Fredy William Bustos

_____

**Matemático** Diego Fernando Ruiz

Fecha de sustentación: Popayán, junio de 2007

# Índice general

# AGRADECIMIENTOS

A Dios, nuestros padres, familiares, docentes y amigos, quienes con su esfuerzo y dedicación hicieron que esta etapa de nuestras vida sea culminada con gran éxito.

# INTRODUCCIÓN

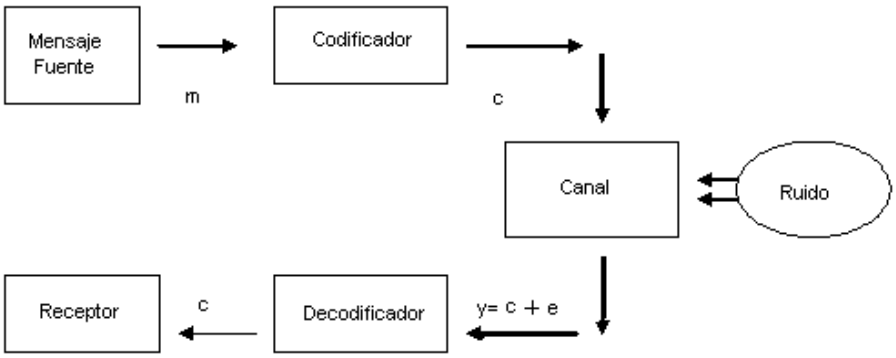
Una de las aplicaciones de los campos finitos es la teoría de codificación, la cual permite la descripción de códigos lineales mediante elementos de un campo $F$ o por medio de polinomios definidos sobre dicho campo. Estas herramientas ofrecen una estructura algebraica bien definida, la cual permite que los códigos lineales sean eficientes. El origen de los códigos correctores de errores se remonta a finales de la década de los 40, cuando los trabajos de Claude Shannon, sobre la teoría de la información abrieron paso a numerosas investigaciones y estudios en el ámbito de las comunicaciones. Inicialmente se abordó como un problema de ingeniería electrónica, con aplicaciones tanto en la transmisión de información como en el almacenamiento de la misma en un soporte digital, siendo su finalidad el preservar la calidad de las comunicaciones frente a la amenaza del ruido, la distorsión o el deterioro del medio. El desarrollo de tales aplicaciones se ha realizado gracias a la utilización de múltiples herramientas matématicas y de la ingeniería.

Uno de los problemas fundamentales en la teoría de codificación es lograr que los errores que ocurren, por ejemplo, debido a canales ruidosos, sean extremadamente improbables. Los métodos para mejorar la confiabilidad de la transmisión dependen de las propiedades de los campo finitos.

Para enviar un mensaje a través de un canal, se debe escribir el mensaje en una forma susceptible de ser enviada por dicho canal. Para ello se traduce el mensaje en un alfabeto $A$ que pueda ser transmitido a través del canal. Siempre se transfiere un número finito de símbolos, lo que impone que el alfabeto sea finito. En este trabajo, el alfabeto corresponde a un campo de dos elementos, por lo cual cada mensaje se puede codificar como una secuencia de ceros y unos. En principio, las palabras escritas en este alfabeto pueden

tener distintas longitudes, sin embargo, para realizar el proceso de corrección y detección de errores es conveniente que todas tengan la misma longitud.

El siguiente gráfico muestra un modelo de un sistema de comunicación. Supongamos que un mensaje de la fuente, denotado por $m$ se desea enviar. Debido a que todo canal está expuesto al ruido, cada mensaje durante la trasmisión puede ser distorcionado, con el riesgo de no ser recuperado en el receptor. La idea básica en teoría algebraica de codificación es transmitir información redundante junto con el mensaje, esto es, se extiende la secuencia de los símbolos mensaje a una secuencia más larga. La redundancia es añadida por el codificador, obteniendo así la palabra código $c$ que es enviada por el canal, donde el ruido representado por un vector error $e$ distorsiona la palabra código, produciendo una palabra recibida $y$. Este vector pasa al decodificador, donde los errores son eliminados, la redundancia es retirada, y se calcula el mensaje original.




En resumen, en este proceso comunicativo intervienen básicamente 5 componentes:

- **Mensaje**: es la información que se quiere comunicar, la cual se trasmite en forma de un vector de $k$ componentes, con elementos en el alfabeto $A$.
- **Fuente**: componente de naturaleza humana o mecánica que determina el tipo de mensaje que se transmitirá y su grado de complejidad.
- **Codificador**: Recurso técnico que transforma el mensaje originado por la fuente, extendiendo cada vector mensaje al agregarle información redundante, convirtiéndolo en una secuencia más larga de $n$ componentes, la cual será enviada por el canal.

- **Canal**: Es el medio por el cual se realiza la transmisión de datos que contienen información, generalmente el canal está expuesto al ruido.

- **Decodificador**: Recurso técnico que transforma las señales recibidas a una forma entendible por el destino, obteniendo el vector mensaje original, eliminando la información distorsionada provocada por la inevitable presencia de interferencia en el canal, con el fin de corregir posibles errores.

# PREFACIO

Entre todos los tipos de códigos, los más estudiados son los códigos lineales, porque su estructura algebraica, permite que sean más faciles de describir, codificar y decodificar que los códigos no lineales. El alfabeto para los códigos lineales es un campo finito, aunque algunas veces, otras estructuras algebraicas como los enteros módulo 4 pueden ser usados para definir códigos que también son llamados lineales. En este texto, estudiamos los códigos lineales, donde el alfabeto es un campo finito de dos elementos, denotado por $F_2$. A lo largo del primer capítulo, nos concentramos en la descripción de la estructura general y las propiedades básicas de los códigos lineales, aunque no consideramos la implementación o aplicación técnica de los códigos; además se plantean las reglas que permiten detectar y corregir errores. En el capítulo 2, se describen algunos tipos de códigos lineales, entre ellos el código Hamming, capaz de corregir un error, con la propiedad de poder extender su tamaño para aumentar la capacidad de corrección y detección de errores. Los siguientes tipos de códigos que se analizan, son los códigos cíclicos. Estos se describen mediante los ideales del anillo cociente $F_2[z]/\langle z^n - 1\rangle$, por lo cual su estructura algebraica está perfectamente definida. Luego, nos referimos a los códigos BCH, los cuales se definen a partir de las raíces del polinomio $q(z) = z^n - 1$, pertenecientes a campos de extensión de $F_2$.

Finalmente, enfatizamos que las aplicaciones se describen unicamente para dar ejemplos del uso real de los códigos, pero no discutimos el diseño experimental.

# Capítulo 1

# CÓDIGOS

Debido a la presencia de señales no deseadas en el proceso de comunicación, cada vector enviado $x \in F_2^n$, puede ser distorsionado convirtiéndose en otro vector $y \in F_2^n$. Para poder corregir estas alteraciones, se debe conocer el número de coordenadas en donde los vectores difieren, por lo cual, se introduce una forma de medida conocida como distancia Hamming.

**Definición 1.** *Sea $F_2$ un campo de dos elementos. Dados dos vectores o palabras $x, y \in F_2^n$, se llama **distancia Hamming** entre ellos, y se denota $d(x, y)$ al número de coordenadas donde ambos vectores no coinciden.*

Al considerar $z \in F_2^n$ usaremos $z_i$ para referirnos a la i-ésima componente de este vector. En cualquier conjunto, en particular en $F_2$, se puede introducir la distancia discreta, donde cualquier par de elementos diferentes distan entre sí la unidad. Por tanto la distancia Hamming en $F_2^n$ se puede expresar así

$$d(x, y) = \sum_{i=0}^{n-1} \delta(x_i, y_i),$$

donde,

$$\delta(x_i, y_i) = \begin{cases} 1 & si\ x_i \neq y_i \\ 0 & \text{en otro caso,} \end{cases}$$

es la métrica discreta en $F_2$.

Para efectos prácticos, se introduce el concepto de peso de un vector, el cual será de gran utilidad para simplificar algunos cálculos.

**Definición 2.** ***El peso Hamming*** *de un vector $x \in F_2^n$, denotado por $w(x)$ es el número de coordenadas no nulas que posee dicho vector.*

Sea

$$\varpi(x_i) = \begin{cases} 1 & si\ x_i \neq 0 \\ 0 & \text{en otro caso,} \end{cases}$$

luego

$$w(x) = \sum_{i=0}^{n-1} \varpi(x_i).$$

**Proposición 1.** *Para cada $x, y, z \in F_2^n$, se cumple*

1. El peso se puede expresar en función de la distancia, es decir $w(x) = d(x, \boldsymbol{0})$.
2. Dados dos vectores $x, y \in F_2^n$ se tiene que $d(x, y) = w(x - y)$.

**Demostración.**

1. La igualdad es inmediata a partir de las definiciones de peso y distancia Hamming.
2. Sea $z = x - y$, donde

$$\varpi(z_i) = \begin{cases} 1 & si\ x_i - y_i \neq 0 \\ 0 & \text{en otro caso} \end{cases} = \begin{cases} 1 & si\ x_i \neq y_i \\ 0 & \text{en otro caso} \end{cases}$$

Por tanto, $w(x - y) = d(x, y)$. $\nabla$

**Teorema 1.** *La distancia hamming define una métrica sobre $F_2^n$. Esto es, para todo $x, y, z \in F_2^n$ se cumple*

1. $d(x, y) \geq 0$ y $d(x, y) = 0$ si y sólo si $x = y$.
2. $d(x, y) = d(y, x)$.

3. $d(x,y) \leq d(x,z) + d(z,y)$.

**Demostración.**

Las dos primeras propiedades son inmediatas a partir de la definición. A continuación se prueba la desigualdad triangular.

Si $x, y \in F_2^n$ entonces $d(x,y) = \sum_{i=0}^{n-1} \delta(x_i, y_i)$, donde $x_i, y_i \in F_2$. Como $\delta$ es la métrica discreta en $F_2$, cumple

$$\delta(x_i, y_i) \leq \delta(x_i, z_i) + \delta(z_i, y_i).$$

Con $z_i \in F_2$, para todo $i = 0 \dots n-1$. Luego

$$\begin{aligned} d(x,y) \leq & \sum_{i=0}^{n-1} (\delta(x_i, z_i) + \delta(x_i, z_i)) \\ = & \sum_{i=0}^{n-1} \delta(x_i, z_i) + \sum_{i=0}^{n-1} \delta(x_i, z_i) \\ = & d(x,z) + d(z,y). \end{aligned}$$

En conclusión, la distancia Hamming es una métrica.$\nabla$

Como el espacio vectorial $F_2^n$, junto con la métrica Hamming es un espacio métrico, tiene sentido considerar el concepto de bola cerrada.

Se denota por $\bar{B}_r(x)$, a la bola cerrada de centro $x$ y de radio $r$, donde $r \leq n$.

$$\bar{B}_r(x) = \{y \in F_2^n : \text{tales que } d(x,y) \leq r\}.$$

En este espacio métrico todas las bolas tienen un número finito de elementos. Es posible contar cuantos elementos tiene.

**Lema 1.** *Si $r \geq 0$ es un entero, el número de elementos de la bola cerrada de radio $r$ centrada en la palabra del código $x$ es*

$$|\bar{B}_r(x)| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{r}.$$

**Demostración.**

Consideremos para $r = 0$. La bola de radio cero tiene un único elemento.

Para $r = 1$ se tiene que la bola contiene su centro y a todos los elementos cuya distancia es exactamente la unidad. Como hay $n$ coordenadas, existen $n$ maneras distintas de elegir una de ellas y considerar el único elemento de $F_2^n$ que difiere del centro sólo en esa coordenada. Hemos visto que la bola de radio uno tiene $\binom{n}{0} + \binom{n}{1}$ elementos.

Para $r = 2$. El número de elementos de esta bola es la suma del resultado anterior y los vectores que se diferencian de $x$ en exactamente dos coordenadas. Para encontrar dichos vectores primero se eligen las dos posiciones, hay $\binom{n}{2}$ posibles formas de escogerlas. Después consideramos el único elemento de $F_2^n$ que difiere de $x$ en esas dos posiciones. Contando, la bola de radio dos tiene

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2},$$

elementos de $F_2^n$.

Para la bola de radio $r$, primero elegimos las $r$ posiciones donde difiere el vector $x$ de $y$, hay $\binom{n}{r}$ posibles formas, después consideramos el único elemento de $F_2^n$ que difiere de $x$ en esas $r$ posiciones, luego se debe sumar todos los elementos que difieren de $x$ en $r$ o menos coordenadas. Por tanto la bola de radio $r$ tiene

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{r}$$

vectores. $\nabla$

En adelante nos referimos a los elementos de $F_2^n$ indistamente como vectores o palabras.

## 1.1. Códigos lineales

**Definición 3.** *Un **código lineal** es un subespacio vectorial $C \subseteq F_2^n$.*

**Definición 4.** *La **distancia mínima** de un código lineal es la menor de las distancias entre dos palabras distintas del código. Esto es,*

$$d_c = \min \{d(x, y) : x \neq y \in C\}.$$

**Definición 5.** *El **peso mínimo** $w_c$ de un código lineal es el menor de los pesos no nulos. Es decir,*

$$w_c = \min \{w(x) : x \in C\}.$$

Un subespacio $C$ de $F_2^n$ de dimensión $k$, contiene $M = 2^k$ vectores de longitud $n$, luego el código lineal $C$ con distancia mínima $d_c$, tiene los siguientes parámetros $[n, k, d_c]$ o también $(n, M, d_c)$.

**Proposición 2.** *Si $C$ es un código lineal entonces $d_c = w_c$.*

**Demostración.**

Mostremos que $d_c \leq w_c$. Sea $w_c$ el peso mínimo de un código lineal $C$, donde $w_c = w(x)$ para algún $x \in C$, como $w(x) = d(x, \boldsymbol{0}) \geq d_c$, luego $w_c \geq d_c$.

Además $d_c \geq w_c$. En efecto, sea $d_c$ la distancia mínima del código lineal, luego existen $x, y \in C$ tales que $d_c = d(x, y) = w(x - y) \geq w_c$, luego $d_c \geq w_c$.

De lo anterior se concluye $d_c = w_c$. $\nabla$

En un código lineal de tamaño $M$, para encontrar la distancia mínima se calcula $\binom{M}{2}$ distancias y de ellas se escoge la menor. Estos cálculos se pueden simplificar gracias a la proposición anterior, dado que, basta con encontrar $M$ pesos.

## 1.2. Corrección y detección de errores

Los códigos se emplean normalmente para transmitir información a través de un canal. Supongamos que el emisor envía una palabra $x \in C$. Si el receptor recibe esa misma palabra no existe ningún problema y puede interpretar correctamente la palabra emitida, pero si el canal deforma la palabra $x$, llegando al receptor la palabra $y$, decimos que se han producido errores en la comunicación. Indicaremos este hecho con la notación $x \longmapsto y$.

**Definición 6.** *Sea $x \longmapsto y$, se denomina **error** al vector $e = x - y$.*

De esta forma, si en la transmisión de una palabra se comete un error solamente, la palabra $y$ difiere de la palabra $x$ en una coordenada, y así sucesivamente.

**Definición 7.** *Dado $x \longmapsto y$, el **número de errores cometidos** es la distancia de $x$ a $y$, es decir, $d(x, y) = \epsilon$.*

**Observación 1.** *Dado $x \longmapsto y$, si $y \in C$ no es posible saber si se han cometido errores en la transmisión.*

Para efectos prácticos, cada código lineal debe permitir analizar la información en el receptor para corregir los errores cometidos en la transmisión, para esto se debe detectar la cantidad de errores. A continuación presentamos una serie de resultados que, dependiendo de las características del código permiten detectar y corregir errores.

**Teorema 2.** *Si $C \subseteq F_2^n$ es un código lineal con distancia mínima $d_c$ entonces $C$ permite detectar $d_c - 1$ errores.*

**Demostración.**
Sea $d_c$ la distancia mínima del código lineal $C$, si al enviar $x$ en $y$ se cometen $d_c - 1$ errores o menos entonces $d(x, y) < d_c$, así que $y \notin C$ y se detectan los errores cometidos, pero si $d(x, y) \geq d_c$ entonces no se puede saber si se han cometido errores. Por lo tanto es posible detectar hasta $d_c - 1$ errores. $\nabla$

**Definición 8.** *Dado un código $C \subseteq F_2^n$, una* ***regla de decisión*** *es una función*

$$f : F_2^n \to (C \cup \{*\}).$$

En una transmisión cuando el canal es afectado por el ruido, podemos recibir cualquier palabra $y \in F_2^n$. Para decodificarla el receptor emplea una regla de decisión, asignándole la palabra $f(y) \in C$ o bien el símbolo $*$ que significa que no puede decodificar la palabra. Decimos que una regla de decisión es capaz de corregir $\epsilon$ errores si decodifica correctamente todas las palabras que tienen un número de errores menor o igual que $\epsilon$.

**Observación 2.** *Una regla de decisión $f$ corrige $\epsilon$ errores si para toda transmisión $x$ en $y$, tal que $d(x, y) \leq \epsilon$ se tiene que $f(y) = x$.*

### 1.2.1. Reglas de decisión

Existen varios algoritmos que permiten corregir errores en la comunicación. A continuación hacemos refencia a algunos de los métodos que generalmente se aplican a cualquier código lineal.

**Definición 9.** *Sea $t$ un número natural,* ***un código lineal*** *$C$* ***corrige*** *$t$* ***errores****, si para todo $y \in F_2^n$ hay a lo más un $x \in C$ tal que $d(x, y) \leq t$.*

- **La regla del vecino más cercano:** Si existe una única palabra del código $x \in C$ a distancia mínima de $y$, entonces $f(y) = x$. Si existen dos o más palabras del código que están a la misma distancia mínima de $y$ entonces $f(y) = *$.

- **La regla de las esferas (bolas) $I$:** Las esferas de radio cero centradas en las palabras del código son disjuntas. Ahora consideremos las esferas de radio unidad, si éstas siguen siendo disjuntas podemos seguir aumentando el radio. Este proceso de aumentar el radio manteniendo disjuntas las esferas debe detenerse en algún momento, dado que el espacio métrico es finito. El mayor radio que permite que las esferas sean disjuntas, se llama radio de empaquetamiento del código.
  Sea $t$ cualquier valor menor o igual que el radio de empaquetamiento, si $y \in \bar{B}_t\ (x)$ entonces $f(y) = x$. Si $y$ no pertenece a ninguna de las esferas entonces $f(y) = *$.

- **La regla de las esferas (bolas) $II$:** Utilizando nuevamente el hecho de que las esferas son disjuntas, para obtener $f(y)$ calculamos la esfera de radio $t$ centrada en $y$. Si dicha esfera contiene a una palabra $x$ del código, entonces $f(y) = x$. Si no contiene a ninguna palabra entonces $f(y) = *$. Si existen dos o más elementos del código en una bola a igual distancia del centro $y$ entonces $f(y) = *$.

**Teorema 3.** *Sea $C$ un código lineal con distancia mínima $d_c$. Si $d_c \geq 2t + 1$ para algún entero positivo $t$ entonces la regla del vecino más cercano permite corregir $t$ errores.*

**Demostración.**

Supongamos que $x \longmapsto y$, y que $d(x, y) \leq t$. Veamos que todas las demás palabras del código están a una distancia superior. Sea $x_1$ otra palabra del código, aplicando la desigualdad triangular tenemos

$$d(x, x_1) \leq d(x, y) + d(y, x_1),$$

luego,

$$d(y, x_1) \geq d(x, x_1) - d(x, y).$$

Como $x$ y $x_1$ pertenecen al código, $d(x, x_1) \geq d_c$ luego,

$$d(y, x_1) \geq d_c - t \geq (2t + 1) - t = t + 1,$$

$$d(y, x_1) \geq t + 1.$$

Por lo tanto $f(y) = x$ y se corrigen los errores cometidos. $\nabla$

**Teorema 4.** *Sea $C$ un código lineal con distancia mínima $d_c \geq 2t + 1$. El método de las esferas $I$ de radio $t$ permite corregir hasta $t$ errores.*

**Demostración.**

Para corregir $t$ errores, se contruyen bolas de radio $t$, centradas en las palabras del código. Supongamos que $u \in \bar{B}_t(x)$ y $u \in \bar{B}_t(x_1)$ con $x, x_1 \in C$, $x \neq x_1$ con $d(x, x_1) \geq d_c$, entonces

$$d(x, x_1) \leq d(x, u) + d(u, x_1) \leq t + t$$

$$d(x, x_1) \leq 2t,$$

lo cual es una contradicción dado que $d(x, x_1) \geq d_c$. Por lo tanto, si $C$ es un código con distancia mínima $d_c$ entonces $C$ puede corregir hasta $t$ errores. $\nabla$

**Observación 3.** *Si $d_c$ es impar, entonces $d_c = 2t + 1$ y permite corregir $t$ errores. Si $d_c$ es par, entonces $d_c = 2t + 2$ y permite corregir $t$ errores. Existe una fórmula que nos proporciona el número $t$ de errores que puede corregir cualquier código, independientemente de si $d_c$ es par o impar*

$$t = \left\lceil \frac{d - 1}{2} \right\rceil,$$

*donde $\lceil\ \rceil$ denota la parte entera techo de un número.*

## 1.3. Cotas del tamaño de los códigos

Los parámetros fundamentales de todo código lineal $C \subseteq F_2^n$ son la longitud $n$, el cardinal $M$ y la distancia mínima $d_c$. El valor de alguno de estos parámetros normalmente limita el posible valor de los otros. Por ejemplo, si $n$ es la longitud del código, no es posible que $d_c$ sea mayor que $n$. De la misma forma no es posible que $M$ supere el valor de $2^n$. Estas limitaciones en los parámetros, dadas en forma de desigualdades, se denominan cotas. Examinaremos ahora algunas de ellas.

**Proposición 3.** ***(Cota Hamming)*** *Sean $t$ un número natural, $C \subseteq F_2^n$ un código lineal con distancia mínima $d_c$ y cardinal $M = |C|$. Si $2t + 1 \leq d_c$ entonces se cumple la desigualdad*

$$M|\bar{B}_t(x)| \leq 2^n.$$

*Es decir,*

$$M \sum_{i=0}^{t} \binom{n}{i} \leq 2^n.$$

**Demostración.**

Consideremos las $M$ bolas de radio $t$ centradas en las palabras del código. Por el teorema 4, estas bolas son todas disjuntas por parejas. El número de vectores en cada una de las $M$ bolas es

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{t},$$

luego, la suma de todos los elementos de las esferas debe ser menor o igual que el número de elementos del espacio vectorial $F_2^n$ que los contiene. Por tanto la suma de todos los elementos es

$$M \sum_{i=0}^{t} \binom{n}{i} \leq 2^n.\nabla$$

Esta desigualdad también se denomina cota de empaquetamiento de las esferas.

**Definición 10.** *Decimos que un código $C \subseteq F_2^n$ es* ***perfecto****, si se alcanza la cota Hamming. Esto es, si*

$$M \sum_{i=0}^{t} \binom{n}{i} = 2^n.$$

**Definición 11.** *Sea $C \subseteq F_2^n$ un código lineal con distancia mínima $d_c$. $C$* ***es maximal*** *si no existe otro código $\acute{C}$, tal que $C \subset \acute{C} \subset F_2^n$ que tenga la misma distancia mínima.*

Como se está trabajando con un número finito de elementos, la existencia de elementos maximales está garantizada.

**Proposición 4.** *Si $C$ es un código perfecto con parámetros $[n, k, d_c]$ entonces es maximal.*

**Demostración.**

Supongamos que $C$ es un código perfecto y no maximal. Existe un código $\acute{C}$ tal que $C \subset \acute{C} \subset F_2^n$ con la misma distancia mínima, tal que $|\acute{C}| > 2^k$. Como $C$ es perfecto cumple que $M|\bar{B}_t(x)| = 2^n$. Además, para ambos códigos se mantiene la relación $d_c \geq 2t+1$, luego $\acute{C}$ cumple que $|\acute{C}||\bar{B}_t(x)| > 2^n$ donde $|\bar{B}_t(x)|$ es el número de vectores en cada una de las esferas de $\acute{C}$, pero esto es una contradicción puesto que las bolas deben estar contenidas en el espacio $F_2^n$. Por tanto si $C$ es perfecto es maximal. $\nabla$

**Observación 4.** *Dados $n$ y $d_c$ es interesante encontrar códigos con el cardinal $M$ lo más grande posible. Denotaremos mediante $A_2(n, d_c)$ a dicho valor de $M$. Este problema se denomina a veces como el problema principal de la teoría combinatoria de códigos. El valor de $M$ está acotado superiormente por tener que cumplir la desigualdad de Hamming.*

**Definición 12.** $A_2(n, d_c) := máx\ \{M$ *tales que un código* $(n, M, d_c)$ *existe* $\}$.

Un código lineal $C$, tal que $M = A(n, d_c)$ es llamado **código lineal óptimo**.

**Proposición 5.** ***(Cota Gilbert-Varshamov)*** *Si $C$ es un código lineal maximal $(n, M, d_c)$ de tamaño $M$ entonces*

$$M \geq \frac{2^n}{|B_{(d_c-1)}(x)|}.$$

**Demostración.**

Si $C$ un código maximal implica que no hay una palabra en $F_2^n$ con distancia $d_c$ o mayor a todas las palabras del código. Es decir, las bolas con centro en las palabras del código son un recubrimiento de $F_2^n$, además la suma de todos los elementos de dichas bolas es $M|B_{(d_c-1)}(x)|$ el cual es mayor o igual que $F_2^n$. $\nabla$

**Observación 5.** *La prueba muestra que un código que tiene al menos $\frac{2^n}{|B_{(d_c-1)}(x)|}$ palabras puede ser construído empezando con una palabra $c_0$ y consucutivamente añadiendo nuevas palabras que tienen distancia al menos $d_c$ a las palabras que han sido escogidas anteriormente, hasta conseguir un código maximal.*

**Proposición 6.** ***(Cota Singleton)*** *Si $C$ es un código lineal $(n, M, d_c)$ entonces*

$$M \leq 2^{n-(d_c-1)}$$

**Demostración.**

Sea $x = (x_1, \ldots, x_{n-(d_c-1)}, \ldots, x_n)$. Consideremos la función

$$\begin{array}{|rl|}\hline \varphi : C \subseteq F_2^n & \to F_2^{n-(d_c-1)} \\ x & \mapsto \varphi(x) := (x_1, \ldots, x_{n-(d_c-1)}) \\ \hline \end{array}$$

Consistente en eliminar las $(d_c - 1)$ últimas coordenadas.

$\varphi$ es función. En efecto, sean $x, z \in F_2^n$ tal $x = z$, en particular por igualdad de vectores se tiene que las $(n - (d_c - 1))$ coordenadas son iguales, es decir, $\varphi(x) = \varphi(z)$.

La aplicación $\varphi$ es inyectiva. Mostremos que el $K(\varphi) = \{\boldsymbol{0}\}$. Supóngase que existe un vector $z \in K(\varphi)$ distinto del vector cero, donde $z = (z_1, \ldots, z_{n-(d_c-1)}, \ldots, z_n)$ con $z_i = 0$ para $1 \leq i \leq n - (d_c - 1)$. Como $w(z) \leq d_c - 1 = w_c - 1$, se presenta una contradicción porque existiría un vector en $C$ con peso menor que el peso mínimo. Por tanto $\varphi$ es inyectiva.

Como los espacios vectoriales son finitos y la función $\varphi$ es inyectiva entonces el número de elementos del código $C$ debe ser menor o igual que el número de elementos del espacio de llegada, es decir,

$$M \leq 2^{n-(d_c-1)}.$$

Lo cual termina la demostración $\nabla$.

**Corolario 1.** *Sea $C$ un código lineal de tipo $[n, k, d_c]$. Entonces $d_c \leq n - k + 1$.*

**Demostración.**

Como $C$ es un código lineal cumple la Cota de Singleton, es decir

$$\begin{aligned} M \leq & 2^{n-d_c+1} \\ 2^k \leq & 2^{n-d_c+1}, \end{aligned}$$

aplicando logaritmo base dos a ambos lados de la igualdad se obtiene, $k \leq n - d_c + 1$. Por tanto $d_c \leq n - k + 1$. $\nabla$

### 1.3.1. Matriz generadora

**Definición 13.** *La aplicación $\varphi : F_2^k \to C \subseteq F_2^n$ se llama* ***función de codificación*** *si $\varphi$ es una transformación lineal.*

**Observación 6.** *Si $m \in F_2^k$ el vector mensaje que se desea codificar, luego $\varphi(m) \in C$ es la palabra codificada, de esta forma el código lineal es precisamente la imagen de la aplicación lineal.*

Sea $m \in F_2^k$ el mensaje generado por la fuente, donde $m = (m_1, m_2, ..., m_k)$ y $\varphi$ una función de codificación tal que $\varphi(m) = x$, con $x \in C \subseteq F_2^n$. Este vector $x$ está formado por los $k$ bits del vector $m$ y $(n-k)$ bits restantes, los cuales se calculan a partir de los bits de mensaje de acuerdo a la regla de codificación establecida. Estos $(n-k)$ bits reciben el nombre de bits de chequeo de paridad. Los códigos de bloque donde los bits de mensaje se transmiten de forma inalterada se denominan **códigos sistemáticos**.
Por ejemplo, supongamos que $m = (m_1, m_2, ..., m_k)$ es un vector mensaje de $k$ componentes. Si un codificador se aplica en esta secuencia de bits, produce una palabra de código de $n$ bits, denotada $x = (x_1, x_2, ..., x_n)$. Si $x$ se cumple

$$x_i = \begin{cases} b_i & \text{i=1,..n-k} \\ m_{i+k-n} & \text{i=n-k+1,..,n} \end{cases}$$

donde $b = (b_1, b_2, ..., b_{n-k})$ es el vector de paridad, entonces el código es sistemático. De acuerdo a nuestro ejemplo, los $(n-k)$ bits a la izquierda de la palabra código son los bits de paridad y los $k$ bits más a la derecha son los bits mensaje, de esta forma nuestra palabra codificada se expresa como el vector

$$x = (b_1, b_2, ..., b_{n-k}, m_{n-k+1}, m_{n-k+2}, ..., m_n).$$

Ahora bien, los $(n-k)$ bits de paridad se calculan a partir de los bits mensaje por medio de combinaciones lineales, denominadas ecuaciones de paridad, por lo cual, podemos escribir

$$b_i = a_{1i}m_1 + a_{2i}m_2 + ... + a_{ki}m_k$$

con $i = 1, \ldots, (n-k)$, donde

$$a_{ji} = \begin{cases} 1 & si\ b_i\ depende\ de\ m_j \\ 0 & en\ otro\ caso. \end{cases}$$

Los coeficientes $a_{ji}$ se eligen de tal manera que los renglones de la matriz generadora (la cual se dá en forma explícita más adelante), sean linealmente independientes y las ecuaciones de paridad sean únicas.

Definamos la matriz $A$ de tamaño $k \times (n-k)$ de coeficientes en $F_2$ de la siguiente manera

$$\mathrm{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1(n-k)} \\ a_{21} & a_{22} & \ldots & a_{2(n-k)} \\ \vdots & \vdots & \ldots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{k(n-k)} \end{pmatrix}$$

Por tanto, obtenemos cada vector $b$ de paridad así, $b = mA$. Por la forma en que está compuesto cada vector $x \in C$ es posible escribir $x = (b, m)$, luego

$$x = \left[ mA \vdots m \right]$$

$$x = m \left[ A \vdots I_k \right].$$

Al denotar por $G = \left[ A \vdots I_k \right]$ se obtiene

$$x = mG.$$

Debido a que $C$ es un subespacio vectorial, se puede extraer de él una base, donde los vectores se colocarán como filas de una matriz.

**Definición 14.** *Llamamos* ***matriz generadora*** *a toda matriz cuyos vectores filas sean linealmente independientes. Decimos que una matriz $G$ de tamaño $k \times n$ genera un código $C$ si los vectores fila de la matriz forman una base del subespacio vectorial $C$.*

El número de columnas de $G$ indica la longitud del código y el número de filas nos informa la dimensión del código.
Sea $\varphi$ una función de codificación, la matriz $M$ asociada a esta aplicación lineal está formada por las imágenes de los vectores de la base, colocados en forma de columna.

**Observación 7.** *En teoría de la codificación se hace un convenio, el cual consiste en multiplicar el vector por la izquierda, esto es, $\varphi(m) = mM$. Si seguimos este convenio los elementos de la base se deben escribir en forma de filas y la matriz $G$ es precisamente la matriz de la aplicación lineal y genera un código $C$, luego $C$ se puede expresar así*

$$C = \left\{ x \in F_2^n : x = mG, m \in F_2^k \right\}.$$

Como cualquier subespacio vectorial de dimensión finita puede tener distintas bases, la matriz generadora $G$ no es única. Por ejemplo, si tomamos una matriz $G$ y permutamos sus filas, la nueva matriz también genera el mismo código. Lo mismo ocurre si a una fila la multiplicamos por un escalar no nulo, o bien sumamos combinaciones lineales de filas. Sabemos, por resultados del álgebra lineal, que aplicando estas operaciones elementales, se pueden trasformar entre sí cualquier par de bases de un subespacio.

**Proposición 7.** *Dos matrices $G$ y $\hat{G}$ generan el mismo código si y sólo si se puede transformar una en la otra, aplicando las siguientes operaciones elementales:*

1. Permutar filas.
2. Multiplicar una fila por un escalar no nulo.
3. Sumar combinaciones lineales de filas a otra fila.

**Definición 15.** *Decimos que un código lineal admite una* ***codificación sistemática*** *si posee alguna matriz generadora de la forma $G = (I_k|A)$, o $G = (A|I_k)$, donde $A$ es de orden $k \times (n-k)$ y contiene la información de los bits de paridad ($I_k$ es debido a que cada palabra codificada tiene como prefijo al vector mensaje).*

**Observación 8.** *En adelante se utilizará la siguiente estructura para la matriz generadora del código $G = (I_k|A)$.*

Si $x = (x_1, \ldots, x_n)$ entonces $x = mG$ por lo cual, $x = (m_1, m_2 ... m_k) * (I_k|A)$ donde

$$x_i = \begin{cases} m_i & i = 1, \ldots, k \\ b_{i-k} = a_{1(i-k)}m_1 + \ldots + a_{k(i-k)}m_k & i = k+1, \ldots, n \end{cases}$$

Luego $x = (m_1, \ldots, m_k, b_1, \ldots, b_{n-k})$. Las últimas $(n-k)$ combinaciones lineales se llaman ecuaciones de paridad.

### 1.3.2. Código Dual

**Definición 16.** *Se llama* ***producto interior*** *o* ***producto escalar*** *de los vectores $x$ e $y$*

$$\langle x, y \rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

*Si se cumple que $\langle x, y \rangle = 0$, decimos que $x$ e $y$ son ortogonales.*

**Definición 17.** *Dado $V \subseteq F_2^n$, se llama* ***conjunto ortogonal*** *de $V$ y se nota $V^\perp$, al conjunto*

$$V^\perp = \{z \in F_2^n : \langle z, w \rangle = 0 \ \forall w \in V\}.$$

**Definición 18.** *Sea $C \subseteq F_2^n$ un código lineal.* ***El código dual*** *de $C$ es el conjunto ortogonal y se denota $C^\perp$. Así,*

$$C^\perp = \{z \in F_2^n : \langle z, x \rangle = 0 \ \forall x \in C\}.$$

Es bueno recordar algunos resultados de espacios vectoriales, los cuales son de gran utilidad al desarrollar la teoría expuesta respecto del código dual.

**Proposición 8.** *Si $C \subseteq F_2^n$ es un código lineal entonces $dim(C) + dim(C^\perp) = dim(F_2^n)$.*

**Proposición 9.** *Si $C$ es un código lineal entonces $(C^\perp)^\perp = C$.*

**Proposición 10.** *$z \in C^\perp$ si y sólo si $z$ es ortogonal a una base de $C$.*

**Demostración.**
$\Rightarrow$) $z$ es ortogonal a una base de $C$. En efecto, si $z \in C^\perp$, entonces es ortogonal a todos los vectores de $C$ y en particular es ortogonal a los elementos de una base.

Sea $B = \{v_1, v_2, \ldots, v_k\}$ una base de $C$ y $z$ ortogonal a cada uno de ellos, es decir, $\langle z, v_i \rangle = 0$ para todo $v_i$.
$\Leftarrow$ Para cada $x \in C$ existen escalares $\alpha_1, \cdots, \alpha_n$, tales que $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$, así

$$\begin{aligned}\langle z, x \rangle =& \langle z, \alpha_1 v_1 + \cdots + \alpha_n v_n \rangle \\ =& \langle z, \alpha_1 v_1 \rangle + \cdots + \langle z, \alpha_n v_n \rangle \\ =& \alpha_1 \langle z, v_1 \rangle + \cdots + \alpha_n \langle z, v_n \rangle \\ =& 0,\end{aligned}$$

esto es, $z$ es ortogonal a cualquier elemento de $C$ y por tanto $z \in C^{\perp}$. $\nabla$

### 1.3.3. Matriz de chequeo de paridad

**Definición 19.** *Dado un código $C$, se llama **matriz de paridad** a una matriz de tamaño $(n-k) \times n$ generadora de su código dual. La denotaremos con la letra $H$.*

Como $H$ genera todos los elementos de $C^{\perp}$, por tanto el código dual se puede expresar de la siguiente forma

$$C^{\perp} = \left\{ u \in F_2^n : u = vH, v \in F_2^{n-k} \right\}.$$

Un vector $w \in F_2^n$ es ortogonal a una base de $C$, escrita como una matriz generadora $G$ si y sólo si $Gw^T = \boldsymbol{0}$ o también, $wG^T = \boldsymbol{0}$.
De esta forma, dado un código $C$, con matriz generadora $G$, podemos afirmar que su código dual es

$$C^{\perp} = \left\{ x \in F_2^n : \ xG^T = \boldsymbol{0} \right\}.$$

La demostración es inmediata a partir de la definición de matriz de chequeo de paridad dado que $G$ es la matriz generadora para $C$.
Como se tiene la igualdad $(C^{\perp})^{\perp} = C$, resulta que $C$ es el espacio nulo de $H$, esto es

$$C = \left\{ x \in F_2^n : \ xH^T = \boldsymbol{0} \right\}.$$

Es decir, los elementos de $C$ son perpendiculares a una base del código dual.

A continuación se presenta una relación importante entre la matriz generadora y la matriz de paridad para un código lineal. Sea $m$ un elemento de $F_2^k$. Sabemos que $x = mG$, donde $x$ es un elemento del código $C$. Si a este elemento le aplicamos $H^T$, debe obtenerse el vector nulo, es decir, $xH^T = \boldsymbol{0}$. Luego

$$\begin{aligned}(mG)H^T &= \boldsymbol{0}\\ m(GH^T) &= \boldsymbol{0} \text{ para todo } m \in F_2^k\\ GH^T &= \boldsymbol{0}\\ HG^T &= \boldsymbol{0}.\end{aligned}$$

En general, dada $G$, encontrar una matriz $H$ que cumpla la condición anterior, nos conduce a un sistema de ecuaciones lineales. Pero si la matriz $G$ está en forma estandar, el cálculo de $H$ es sencillo.

**Proposición 11.** *Si $G$ está escrita en forma estandard $(I_k, A)$, entonces la matriz de paridad es $H = (-A^T, I_{(n-k)})$.*

**Demostración.**

Supongase que la matriz $H$ es del tamaño y del rango adecuados, así $HG^T = 0$. En efecto,

$$\begin{aligned}HG^T &= (-A^T|I)(I|A)^T\\ &= (-A^T|I)(I|A^T)\\ &= -A^T + A^T\\ &= \boldsymbol{0},\end{aligned}$$

utilizando la multiplicación matricial por bloques. $\nabla$

**Definición 20.** *Una matriz de paridad $H$ está en forma estandar si $H = (-A^T, I_{(n-k)})$.*

**Observación 9.**

- *La matriz generadora $G$ se usa en la operación de codificación en el transmisor.*
- *La matriz $H$ de verificación de paridad se usa en la operación de decodificación en el receptor.*

Una de las principales ventajas que tiene la matriz de paridad sobre la matriz generadora es que permite calcular la distancia mínima, como lo garantiza el siguiente teorema.

**Teorema 5.** *Sea $C$ un código lineal con matriz de chequeo de paridad $H$. Si $d_c$ es la distancia mínima de $C$ entonces $d_c$ es el menor número de columnas linealmente dependientes de $H$.*

**Demostración.**

Sea $x \in C$, con matriz de chequeo de paridad

$$\mathrm{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \cdots & h_{(n-k)n} \end{pmatrix}$$

Denotemos mediante $h_i$, $1 \leq i \leq (n-k)$, las columnas de la matriz $H^T$. Es claro que $xH^T = \boldsymbol{0}$. Luego

$$xH^T = (xh_1, xh_2, \ldots, xh_(n-k)) = (0, 0, \ldots, 0).$$

Por lo cual

$$\begin{aligned} xh_1 &= x_1h_{11} + x_2h_{12} + \ldots + x_nh_{1n} = 0 \\ xh_2 &= x_1h_{21} + x_2h_{22} + \ldots + x_nh_{2n} = 0 \\ \vdots &\quad \vdots \\ xh_{(n-k)} &= x_1h_{(n-k)1} + x_2h_{(n-k)2} + \ldots + x_nh_{(n-k)n} = 0 \end{aligned}$$

Sumando y agrupando se obtiene

$$x_1 \begin{pmatrix} h_{11} \\ h_{21} \\ \vdots \\ h_{(n-k)1} \end{pmatrix} + x_2 \begin{pmatrix} h_{12} \\ h_{22} \\ \vdots \\ h_{(n-k)2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} h_{1n} \\ h_{2n} \\ \vdots \\ h_{(n-k)n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Se denota $\hat{h}_j$ con $1 \leq j \leq n$ a las filas de $H^T$. Claramente se verifica que

$$x_1\hat{h}_1 + x_2\hat{h}_2 + x_3\hat{h}_3 + \ldots + x_n\hat{h}_n = 0.$$

Dado que $C$ es un código lineal con distancia mínima $d_c$, existe un vector $x \in C$ tal que $w(x) = d_c = w_c$, es decir, el vector $x$ tiene $d_c$ componentes no cero, los cuales son los coeficientes de la combinación lineal de las filas de $H^T$, entonces indezando nuevamente las componentes no cero del vector $x$, obtenemos la combinación lineal

$$x_1\hat{h}_1 + x_2\hat{h}_2 + \ldots + x_{d_c}\hat{h}_{d_c} = 0.$$

Por tanto, hay $d_c$ columnas linalmente dependientes.

Ahora se prueba que $d_c$ es el menor número de columnas linealmente dependientes. Supongamos que hay $d$ columnas de $H$ linealmente dependientes con $d < d_c$. Entonces existen $\alpha_1, \alpha_2, \ldots, \alpha_d$ elementos no todos cero en el campo $F_2$, tales que

$$\alpha_1\hat{h}_1 + \alpha_2\hat{h}_2 + \ldots + \alpha_d\hat{h}_d = 0.$$

Con estos coeficientes y aumentando ceros ($\alpha_i = 0$), se construye un vector $z \in F_2^n$ que cumple:

$$\alpha_1\hat{h}_1 + \alpha_2\hat{h}_2 + \ldots + \alpha_n\hat{h}_n = 0.$$

así, $zH^T = 0$ por definición de $C$, luego $z \in C$.

Como $w(z) = d < d_c = w_c$ esto contradice la elección de $w_c$, dado que no puede existir un vector en $C$ con menor peso que el mínimo. En conclusión, $d_c$ es el menor número de columnas linealmente dependientes de $H$. $\nabla$

**Corolario 2.** *Sea $C$ un código lineal $[n, k, d_c]$ con matriz de chequeo de paridad $H$. Si $d$ es el menor número de columnas linealmente dependientes de $H$ entonces $d = d_c$.*

**Demostración.**

Por el teorema anterior, $d_c$ es el menor número de columnas linealmente dependientes de $H$, luego $d_c = d$. $\nabla$

**Corolario 3.** *Si $C$ es un código lineal con distancia mínima $d_c$ y matriz de paridad $H$, entonces $d_c - 1$ es el mayor número de columnas de $H$ linealmente independientes.*

**Demostración.**

Sea $d_c$ la distancia mínima de $C$, por el teorema 5, $d_c$ es el menor número de columnas de $H$ linealmente dependientes, luego hay $d_c - 1$ columnas linealmente independientes. $\nabla$

### 1.3.4. Cálculo del error

Debido a la presencia de ruido en el canal durante el envío de datos, al receptor llega una palabra $y$ que es el resultado de sumar, al vector original $x$, un vector $e$, el cual es el error generado por la contaminación del medio de transmisión.
Si $y = x + e$, donde $e = (e_1, e_2, \ldots, e_n)$ entonces

$$e_i = \begin{cases} 1 & \text{si un error ha ocurrido en la i-ésima posición.} \\ 0 & \text{en otro caso.} \end{cases}$$

A continuación se describe un algoritmo de decodificación para códigos lineales. Sea $C$ un código lineal $[n, k, d_c]$ sobre $F_2$. El anillo cociente $F_2^n/C$ consiste de todas las clases laterales

$$a + C = \{a + c : c \in C\}.$$

Cada clase lateral contiene $2^k$ elementos, donde $F_2^n$ puede verse como

$$F_2^n = (a^0 + C) \cup (a^1 + C) \cup (a^2 + C) \cup \ldots \cup (a^{2^{n-k}-1} + C).$$

El receptor decodifica el vector enviado $x$ a partir del vector recibido $y$, este procedimiento puede facilitarse haciendo uso de las clases laterales $F_2^n/C$.

**Definición 21.** *Sea $C \subseteq F_2^n$ un código lineal de tipo $[n, k, d_c]$ y el anillo cociente $F_2^n/C$. Un elemento de peso mínimo en la clase $a + C$ es llamado un representante de la clase; si varios vectores en $a + C$ tienen peso mínimo entonces se escoge uno de ellos como representante de dicha clase.*

Sean $a^1, a^2, \cdots, a^{2^{n-k}-1}$ los representantes de las clases laterales diferentes a la clase $C$ y sean, $x^1, x^2, x^3, \cdots x^{2^k}$ todas las palabras código en $C$, donde $x^1 = 0$. Consideremos el siguiente arreglo

$$\begin{cases} 0 + x^1 & 0 + x^2 & \ldots & 0 + x^{2^k} \\ a^1 + x^1 & a^1 + x^2 & \ldots & a^1 + x^{2^k} \\ \vdots & \vdots & & \vdots \\ a^{2^{n-k}-1} + x^1 & a^{2^{n-k}-1} + x^2 & \ldots & a^{2^{n-k}-1} + x^{2^k}. \end{cases}$$

Si una palabra $y \in F_2^n$ es recibida, con $y = a^i + x^j$, utilizando las reglas de decisión, el decodificador decide que el vector error $e$ corresponde al representante de la clase lateral $a^i$ (el de menor peso), el cual, garantiza la menor distancia entre $x$ e $y$, siempre que, en esta clase se encuentre un solo vector de peso mínimo; de esta forma se decodifica $y$ como la palabra $x = y - e = x^j$.

**Definición 22.** *Dada la clase $a+C$, llamamos **líder de la clase** al único vector $e \in a+C$ de peso mínimo.*

**Observación 10.** *Es claro que en cada clase debe existir un vector de peso mínimo. Si existen varios vectores de peso mínimo, decimos que esa clase no tiene lider. Sin embargo el siguiente lema garantiza la existencia del lider en los casos que nos interesa.*

**Teorema 6.** *Sea $C$ un código de distancia mínima $d_c \geq 2t + 1$. Si $d(x, y) \leq t$ donde $y \in (a + C)$ entonces la clase $a + C$ tiene lider.*

**Demostración.**
Sea $x$ un elemento del código que cumple $d(x, y) \leq t$. Construimos el vector $e = y - x$ que necesariamente pertenece a $a + C$. Entonces el peso de $e$ es menor que $t$ puesto que $d(x, y) = w(y - x) = w(e)$. Si $e^*$ es otro vector de la clase, de peso menor o igual que $t$, se cumple: $e = a + x_i$ y $e^* = a + x_j$, entonces $e - e^* = x_i + x_j$, es decir, $(e - e^*) \in C$, pero $w(e - e^*) \leq w(e) + w(e^*) \leq 2t < d_c$, lo cual contradice la definición de distancia mínima.

### 1.3.5. Síndrome

El receptor tiene la tarea de decodificar el vector de código $x$, a partir del vector recibido $y$. El algoritmo que se usa comunmente para efectuar esta operación de decodificación empieza con el cálculo de un vector de tamaño $1 \times (n - k)$ llamado vector síndrome del error o simplemente vector de diagnóstico. La importancia de este síndrome radica en que depende solamente del errror.

**Definición 23.** *Sean $C \subseteq F_2^n$ un código lineal, $H$ una matriz de chequeo de paridad y el vector $y \in F_2^n$. Se nota por $s_y$ **al síndrome de** $y$, el cual corresponde al vector $yH^T$.*

**Proposición 12.** *El vector síndrome cumple las siguientes propiedades*

1. El síndrome de un elemento es nulo si y sólo si el elemento pertenece al código.
2. $s(y) = s(z)$ si y sólo si $y + C = z + C$.
3. El síndrome depende únicamente del error y no de la palabra de código transmitida.

**Demostración.**

1. Por definición de $C$ se tiene que $s_z = 0$ si sólo si $zH^T = \boldsymbol{0}$, para todo $z \in C$.
2. $s_y = s_z$ si y sólo si $zH^T = yH^T$ si y sólo si $(z-y)H^T = \boldsymbol{0}$ si y sólo si $(z-y) \in C$ si y sólo si $z + C = y + C$.
3. Sea $y \in F_2^n$ una palabra recibida, luego

$$\begin{aligned} s_y =& yH^T \\ =& (x+e)H^T \\ =& xH^T + eH^T. \end{aligned}$$

   Como $x \in C$, se tiene $xH^T = \boldsymbol{0}$, además $s_e = eH^T$, luego $s_y = s_e$. $\nabla$

La igualdad anterior indica que la palabra recibida y el vector error pertenecen a la misma clase lateral. Sea $y = x + \tilde{e}$ donde $\tilde{e}$ es un vector error, $y \in a + C$ y esta clase contiene todas las expresiones de la forma $\tilde{e} = y - x$. Si escogemos un vector error $e$ de menor peso en esta clase y decodificamos $y$ como $x = y - e$, entonces $x$ será un vector en $C$ a distancia mínima de $y$. Pueden haber varios vectores de peso menor, en tal caso escogemos uno de ellos. Por la definición de líder se tiene que el vector error $e$ coincide con él.

**La regla de decisión del líder para decodificar sigue los siguientes pasos**

1. Dada la transmisión $x \longmapsto y$, calculamos el síndrome de $y$. Si es nulo, entonces se hace $f(y) = x$.

2. Si el síndrome es no nulo y su clase asociada tiene lider $e$, entonces se hace

   $f(y) = y - e$.

3. Si el síndrome es no nulo y la clase asociada no tiene lider, entonces $f(y)$ no decide.

El lema anterior nos dice que si $d(x, y) \leq t$, la clase asociada tiene líder y por tanto esta regla de decisión es capaz de corregir hasta $t$ errores. Es fácil ver que esta regla coincide con la del vecino más cercano, a pesar de que el método de construcción sea distinto.

**Teorema 7.** *Existe una correspondencia biunívoca entre el conjunto de clases de equivalencia de $F_2^n$ módulo $C$ y el conjunto de síndromes.*

**Demostración.**

Sea $S$ el conjunto formado por los síndromes, consideremos la aplicación

$$\boxed{\begin{array}{ll} \varphi: & \mathrm{F}_2^n/C \to S \\ & \mathrm{a{+}C} \mapsto \varphi(a + C) = s_a \end{array}}$$

Probemos que $\varphi$ es función. Sean $z + C$, $y + C$ en $F_2^n/C$, con $z + C = y + C$. Ahora, $\varphi(z + C) = \varphi(y + C)$ si y sólo si $s_z = s_y$.

$\varphi$ es inyectiva. Sean $z + C$, $y + C$ en $F_2^n/C$, supongamos que $\varphi(z + C) = \varphi(y + C)$, esto es, $s_z = s_y$ si y sólo si $z + C = y + C$, porque el conjunto de clases laterales forman una partición de $F_2^n$.

Finalmente, $\varphi$ es sobreyectiva. En efecto, para cada $s_y \in S$, con $y \in F_2^n/C$, existe una única clase lateral $a + C$, tal que $y \in a + C$.

Esta función permite asignar a cada clase lateral un único vector síndrome. $\nabla$

#### 1.3.5.1. Algoritmo de decodificación

Sea $\mathtt{M} = \{m_1, m_2, \ldots, m_{2^k}\}$, un conjunto de vectores mensajes a codificar y el conjunto $C = \left\{x^1 = 0, x^2, x^3, \cdots, x^{2^k}\right\}$, un código lineal $[n, k, d_c]$, con matriz de paridad $H$.

En el anillo cociente $F_2^n/C$, sean $a^1, a^2, \cdots, a^{2^{n-k}-1}$ los representantes de las clases laterales diferentes a la clase $C$. Las palabras código forman la clase del cero, para hallar las otras clases del conjunto $F_2^n/C$, tomamos un elemento de $F_2^n$ que no esté en el código y

sumamos módulo 2 con cada elemento del código. Luego de formar las clases laterales, se toma un representante de cada una de ellas (el de menor peso) y se calcula el síndrome de cada representante, de esta forma se encuentra el síndrome de cada clase. Consideremos el siguiente arreglo

| | | | |
|---|---|---|---|
| $x^1$ | $x^2$ | $\dots x^{2^k}$ | $\rightarrow s_0$ |
| $a^1 + x^1$ | $a^1 + x^2$ | $\dots\ a^1 + x^{2^k}$ | $\rightarrow s_{a^1}$ |
| $\vdots$ | $\vdots$ | | $\vdots$ |
| $a^{2^{n-k}-1} + x^1$ | $a^{2^{n-k}-1} + x^2$ | $\dots\ a^{2^{n-k}-1} + x^{2^k}$ | $\rightarrow s_{a^{2^{n-k}-1}}$ |

Si $y$ es la palabra recibida, en el arreglo se busca la clase a la que pertenece $y$; para arreglos grandes este procedimiento consume demasiado tiempo, para evitar estas dificultades se hace uso del vector síndrome $s_y$, conociendo que $s_y = yH^T = s_{a^i}$, el cual informa que $y$ pertenece a la clase que tiene dicho síndrome. Como el vector error $e$ es igual al lider de la clase $a^j$ el cual tiene el mismo síndrome de $y$, se puede encontrar la palabra código original enviada, al decodificar $y$, así

$$y = x + e$$

$$x = y - e.$$

Para encontrar el mensage original $m$, dado $G = (I_k, A)$ basta con tomar las primeras $k$ componentes del vector código $x$.

**Ejemplo 1.** *Sea $C$ un código lineal binario con parámetros $[4, 2, d_c]$, con matriz generadora $G$ y matriz de chequeo de paridad $H$*

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

**Conjunto de palabras mensaje** $\mathtt{M} = \{00, 10, 01, 11\}$

**Código** $C = \{0000, 1010, 0111, 1101\}$.

Se forman las clases laterales, se toma un representante de cada clase (el de menor peso) y se calcula el síndrome de cada representante, de esta forma se encuentra el síndrome de cada clase. El correspondiente arreglo de clases laterales es

| | | | | |
|---|---|---|---|---|
| 0000 | 1010 | 0111 | 1101 | $\rightarrow \binom{0}{0}$ |
| 1000 | 0010 | 1111 | 0101 | $\rightarrow \binom{1}{0}$ |
| 0100 | 1110 | 0011 | 1001 | $\rightarrow \binom{1}{1}$ |
| 0001 | 1011 | 0110 | 1100 | $\rightarrow \binom{0}{1}$ |

Si la palabra $y = 1110$ es recibida, en el arreglo se busca la clase a la que pertenece $y$, conociendo que $s_y = yH^T = \binom{1}{1}$, entonces $y$ pertenece a la clase que tiene dicho síndrome. Como el vector error $e$ es igual al lider de la clase, el cual es 0100, se puede encontrar la palabra código original enviada, al decodificar $y$, así

$$\begin{aligned} y =& x + e \\ x =& y - e && \textit{reemplazando}, \\ x =& 1110 - 0100 \\ x =& 1010 && \textit{además}, \\ m =& 10, \end{aligned}$$

donde $m$ es el mensage original sin codificar.

En códigos lineales con longitud $n$ grande, el proceso para encontrar los representantes de peso mínimo de las clases laterales se hace extenso. Para sortear tales dificultades se hace uso del siguiente resultado.

**Teorema 8.** *Sea $C \subseteq F_2^n$ un código lineal $[n, k, d_c]$, con matriz de chequeo de paridad $H$, el síndrome es la suma de las columnas de $H$ que corresponden a posiciones donde ha ocurrido error.*

**Demostración.**

Sea $y \in F_2^n$ el vector recibido, luego $y = x + e$ con $x \in C$. Sabemos que $s_y = yH^T = eH^T$. Se toma las coordenadas no nulas en el vector error $e$, indezando nuevamente se tiene que dichas coordenadas son $e_1, e_2, \ldots, e_t$, donde $t$ toma valores entre 1 y $n$, luego al realizar

el producto $eH^t$ se tiene $s_y = h_1 + \ldots + h_i + \cdots + h_t$ donde $h_i$ corresponde a la i-ésima columna de $H$. $\nabla$

## 1.4. Códigos con distancia de separación máxima

Existen códigos que son más eficientes que otros, entre ellos los llamados MDS, los cuales tienen la característica de que su distancia mínima $d_c$ es lo mas grande posible; esto permite que tales códigos tengan la mayor capacidad de corregir errores.

**Corolario 4.** *Si $C$ es un código lineal de tipo $[n, k, d_c]$ entonces $d_c \leq n - k + 1$.*

**Demostración.**
Como $C$ es lineal, satisface la cota Singleton, luego $2^k \leq 2^{n-d_c+1}$, pero la función exponencial es creciente, por tanto $k \leq n - d_c + 1$, esto es, $d_c \leq n - k + 1$.

**Definición 24.** *Un código lineal $C$ con* ***distancia de separación máxima*** *es aquel que cumple $d_c = n - k + 1$. También de denomina código* MDS.

**Proposición 13.** *Un código lineal $C$ con parámetros $[n, k, d_c]$ es* MDS *si y sólo si un conjunto de $n - k$ columnas de una matriz de chequeo de paridad $H$ de $C$ forma un conjunto linealmente independiente de vectores de $F_2^{n-k}$.*

**Demostración.**
$\Rightarrow$) Sea $H$ la matriz de chequeo de paridad para $C$, entonces por corolario 3, como $C$ tiene distancia mínima $d_c$ entonces todo conjunto de $d_c - 1$ columnas de $H$ son linealmente independientes. Debido a que $C$ es MDS cumple que $d_c - 1 = n - k$.
$\Leftarrow$) Si $H$ tiene $n - k$ columnas linealmente independientes, entonces $n - k \leq d_c - 1$ dado que $d_c - 1$ es el mayo número de columnas linealmente independientes de $H$. Como $C$ es lineal satisface el corolario 4, es decir, $d_c - 1 \leq n - k$, luego se cumple la igualdad $d_c = n - k + 1$. Por lo tanto $C$ es un código MDS. $\nabla$

**Proposición 14.** *Si $C$ es un código* MDS *con parámetros $[n, k, d_c]$ entonces $C^\perp$ también es* MDS *con parámetros $[n, n - k, d_{c^\perp}]$.*

**Demostración.**

Se debe probar que la distancia mínima de $C^{\perp}$ satisface la igualdad $k+1 = d_{c^{\perp}}$. Sea $H$ la matriz de paridad del código lineal $C$ de dimensión k. Como el código es MDS, cualquier conjunto de $r = n - k$ columnas de $H$ es siempre linealmente independiente. Como $H$ es la matriz generadora del dual, todo elemento $v$ de este, se puede escribir en la forma

$$v = uH = (\langle h_1, u\rangle, \langle h_2, u\rangle, \ldots, \langle h_n, u\rangle),$$

donde $h_i$ denota la i-ésima columna de $H$ y $u$ es un vector de longitud $r$. El número máximo de elementos nulos de este vector es $r - 1$. Veamos que un número mayor nos lleva a contradicción. Supongamos que hay $r$ coordenadas nulas de $v$, entonces existen $r$ columnas que cumplen la condición $\langle h_i, u\rangle = 0$. Las $r$ columnas pertenecen al ortogonal de $u$, cuya dimensión es menor o igual que $r - 1$, lo cual es una contradicción dado que las $r$ columnas son linealmente independientes. Luego, el número de coordenadas no nulas de $v$ debe ser menor o igual que $r - 1$, entonces $w(v)$ debe ser mayor o igual que $n - (r - 1) = n - (n - k - 1) = k + 1$. Pero, $w_{c^{\perp}} = d_{c^{\perp}}$, se tiene que $d_{c^{\perp}} \geq k + 1$.
Como $C^{\perp}$ es un código lineal, por corolario 1 se cumple que $d_{c^{\perp}} \leq k+1$. Así que $k+1 = d_{c^{\perp}}$.
Por tanto $C^{\perp}$ es un código MDS. $\nabla$

**Corolario 5.** *$C$ es un código lineal* MDS *si y sólo si cada conjunto de $k$ columnas de la matriz generadora $G$ de $C$ es linealmente independiente.*

**Demostración.**

$\Rightarrow$) La matriz $G$ es la matriz de chequeo de paridad para el $C^{\perp}$, luego hay $d_{c^{\perp}} - 1 = k$ columnas linealmente independientes de la matriz $G$.
$\Leftarrow$) Supongamos que la matriz $G$ tiene $k$ columnas linealmente independientes, luego $k \leq d_{c^{\perp}} - 1$ y como $C^{\perp}$ es lineal, se tiene por el corolario 1, que $d_c - 1 \leq n - (n - k) = k$.
Por tanto se cumple $d_{c^{\perp}} = k + 1 = n - (n - k) + 1$. Es decir, $C^{\perp}$ es un código MDS. $\nabla$

#### 1.4.0.2. Matriz de Vandermonde

Sea un campo $F$, $\alpha_1, \alpha_2, \ldots, \alpha_n$ elementos distintos en $F$. Decimos que una matriz $D_n$ de tamaño $n \times n$ es de Vandermonde si $D_n$ tiene una de las siguientes formas

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ {\alpha_1}^2 & {\alpha_2}^2 & \cdots & {\alpha_n}^2 \\ \vdots & \vdots & & \vdots \\ {\alpha_1}^{n-1} & {\alpha_2}^{n-1} & \cdots & {\alpha_n}^{n-1} \end{pmatrix} \text{o} \begin{pmatrix} 1 & \alpha_1 & {\alpha_1}^2 & \cdots & {\alpha_1}^{n-1} \\ 1 & \alpha_2 & {\alpha_2}^2 & \cdots & {\alpha_2}^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & \alpha_n & {\alpha_n}^2 & \cdots & {\alpha_n}^{n-1} \end{pmatrix},$$

donde su determinante es no nulo.

**Teorema 9.** *Sean $\alpha_1, \alpha_2, \ldots, \alpha_n$ elementos distintos de un campo $F$ y sea $d < n$. El código $C \subseteq F^n$ que tiene por matriz de paridad*

$$H=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ {\alpha_1}^2 & {\alpha_2}^2 & \cdots & {\alpha_n}^2 \\ \vdots & \vdots & & \vdots \\ {\alpha_1}^{d-2} & {\alpha_2}^{d-2} & \cdots & {\alpha_n}^{d-2} \end{pmatrix}$$

*es un código* MDS*, con parámetros $[n, n-(d_c-1), d=d_c]$.*

**Demostración.**

Como $C \subseteq F^n$, luego los vectores $x \in C$ tienen longitud $n$.

Para la dimensión de $C$, recuerde que $R_H + dim(C) = n$, donde $R_H$ denota el rango de $H$, además $C$ corresponde al espacio nulo de la matriz $H$, dado que

$$C = \left\{ x \in F_2^n : \ xH^T = 0 \right\}.$$

Sea $R$ el espacio generado por los renglones de $H$, luego $dim(R) = R_H$, como $H$ es la matriz generadora del código dual, por definición sus filas forman un conjunto linealmente independiente, por tanto las $d-1$ filas son una base para $R$, así $dim(R) = d-1 = R_H$. En conclusión $dim(C) = n-(d-1)$.

Sea $C_H$ el espacio generado por las columnas de $H$, como $dim(C_H) = dim(R) = R_H$, entonces a lo más, hay $d-1$ columnas linealmente independientes las cuales forman una

base de $C_H$. Además $d$ es el menor número de columnas linealmente dependientes, por corolario 2, $d = d_c$.

Finalmente se muestra que es un código MDS. En efecto,

$$\begin{aligned} d =& d_c \\ =& n - n + 1 - 1 + d \\ =& n - (n - (d - 1)) + 1 \\ =& n - k + 1. \end{aligned}$$

Lo cual concluye la demostración. $\nabla$

## 1.5. Código Extendido

Si a cada palabra del código lineal se le añade un nuevo dígito, de tal forma que al sumar todas las coordenadas, incluyendo la nueva componente la suma se hace cero, se obtiene un nuevo código, llamado código extendido. Este proceso se realiza cuando se desea aumentar la capacidad de un código para corregir errores.

**Códigos con bit de paridad**

- **Paridad par:** Dado $x \in F_2^n$, $x = (x_1, x_2, \ldots, x_n)$ construimos $\hat{x} \in F_2^{n+1}$, donde $\hat{x} = (x_1, x_2, \ldots, x_n, x_{n+1})$ añadiendo un 0 si $w(x)$ es par y un 1 si $w(x)$ es impar, donde $x_{n+1} = \sum_{i=1}^{n} x_i (\text{mód} 2)$. Se tiene que $C$ es de tipo $[n+1, n-r, 2i]$, con $i$ un número natural. Según la construcción, el peso de todos los vectores es par. Como el código es lineal, se cumple que $d_c = w_c$, por tanto $d_c$ es par.

- **Paridad impar:** Dado $x \in F_2^n$, $x = (x_1, x_2, \ldots, x_n)$ construimos $\hat{x} \in F_2^{n+1}$, con $\hat{x} = (x_1, x_2, \ldots, x_n, x_{n+1})$ añadiendo un 0 si $w(x)$ es impar y un 1 si $w(x)$ es par, donde $x_{n+1} = \sum_{i=1}^{n} x_i (\text{mód} 2)$. Se tiene que $C$ es de tipo $[n+1, n-r, 2i+1]$, con $i$ en los naturales. Según la construcción, el peso de todos los vectores es impar. Como el código es lineal, se cumple que $d_c = w_c$, por tanto $d_c$ es impar.

Lo anterior establece una biyección de $F_2^n$ con un subconjunto $C$ de $F_2^{n+1}$ que será nuestro nuevo código.

**Definición 25.**

Si $C \subseteq F_2^n$ es un código de longitud $n$, se define el código extendido $C_E$ como

$$C_E := \left\{ (x_1, x_2, \ldots, x_n, x_{n+1}) : (x_1, x_2, \ldots, x_n) \in C,\ x_{n+1} = \sum_{i=1}^{n} x_i m\acute{o}d(2) \right\}.$$

Sea $C$ es un código lineal con matriz de paridad $H$ entonces $C_E$ tiene matriz de chequeo de paridad $H_E$, para obtener la matriz generadora $G_E$ se resuelve el sistema $HG^T = \boldsymbol{0}$.

# Capítulo 2

# CÓDIGOS LINEALES

## 2.1. Código Hamming

Hasta aquí se ha presentado la estructura general de los códigos lineales. Ahora se analizará algunos tipos de códigos particulares. En este capítulo se inicia con el código Hamming, el cual tiene la capacidad de detectar dos errores y corregir uno. En general los códigos lineales que tienen esta característica se llaman códigos de corrección simples. Si se desea ampliar la capacidad de corrección de errores se recurre al código Hamming extendido. Estos algoritmos fueron construidos por Richard Hamming, el cual fue uno de los pioneros en la teoría de la codificación.

**Definición 26.** *Un código binario $C$ de longitud $n = 2^r - 1$, con $n - k = r \geq 2$ y matriz $H$ de chequeo de paridad de tamaño $r \times (2^r - 1)$, se llama* ***código Hamming****, si las columnas de $H$ son las representaciones binarias de los enteros $1, 2, \ldots, (2^r - 1)$, donde cada par de columnas son linealmente independientes.*

Dado un natural $r$, considérese todos los números binarios no nulos de $r$ cifras (en total hay $n = 2^r - 1$); con estos se construye la matriz de paridad $H$ ubicándolos como columnas en orden creciente, donde el total de filas de $H$ es $r$. Esta matriz es de rango $r$, porque la base canónica de $F_2^n$ está incluida dentro de ella y se satisface para $C_H$ (el espacio de las

columnas de $H$) la siguiente igualdad

$$dim C_H = dim(imagen H) = r.$$

Se denota el código Hamming como

$$\text{HAM}(r) = \{x \in F_2^n : \ xH^T = 0\}.$$

Tenemos que HAM$(r)$ es el espacio nulo de la matriz $H$, para el cual se cumple que $r + dim\text{HAM}(r) = n$, donde $n$ es el número de columnas. Por tanto el código Hamming es un subespacio vectorial de dimensión $k = (n - r) = 2^r - 1 - r$.
Los códigos Hamming tienen la característica de que su distancia mínima $d_c$ es igual a 3, como lo afirma el siguiente teorema.

**Teorema 10.** *Un código Hamming tiene peso mínimo igual a* 3.

**Demostración**.
Mostremos que no existen elementos de peso 1 y 2. Si $x \in$ HAM$(r)$ donde $w(x) = 1$, entonces tal vector pertenece a la base canónica, el cual se denota mediante $x = e_i$, con $i = 1, \ldots, r$. Al actuar $H$ sobre un vector $e_i$, se tiene $e_i H^T = h_i$, donde $h_i$ es una columna de $H$ que por definición es no nula. Si $w(x) = 1$ entonces $xH^T \neq \boldsymbol{0}$ y por tanto no pertenece al subespacio HAM$(r)$.
Si $w(x) = 2$, al actuar $H^T$ sobre él se obtiene la suma de dos columnas de la matriz, es decir $xH^T = h_i + h_j$, como son linealmente independientes, necesariamente $h_i + h_j \neq \boldsymbol{0}$, así $xH^T \neq \boldsymbol{0}$. Por tanto $x$ no pertenece al subespacio.
Mostremos que existe un vector $x \in$ HAM$(r)$ tal que $w(x) = 3$. Sea el vector $x \in F_2^n$, donde solamente sus tres primeras componentes son diferentes de cero. Al multiplicar matricialmente se tiene $xH^T = h_1 + h_2 + h_3$. Al sumar las dos primeras columnas se obtiene la tercera, dado que son las representaciones binarias de los enteros $1, 2, 3$, es decir $h_1 + h_2 = h_3$, entonces al sumar módulo 2 se cumple que $xH^T = \boldsymbol{0}$. Por tanto $x \in$ HAM$(r)$. $\nabla$
Como el código Hamming es lineal, cumple que su peso mínimo es igual a su distancia mínima, es decir $w_c = d_c$, luego la distancia mínima es igual a 3.

**Corolario 6.** *El código* HAM$(r)$ *con parámetros* $[n, n-r, d_c]$ *detecta dos errores y corrige uno.*

**Demostración.**

Como los códigos de Hamming tienen $d_c = 3$, permiten detectar $d_c - 1$ errores y corrige hasta $t$ errores, siempre que $2t + 1 \leq d_c$ para algún $t$. Como $d_c = 3$ se tiene que $t = 1$. Por tanto el código hamming es un corrector simple de error. $\nabla$

Si al enviar $x \to y$ se cometen dos errores, entonces $y = x + e_i + e_j$. Si hacemos actuar la matriz $H$ sobre el vector $y$ se tiene

$$\begin{aligned} s_y = yH^T &= (x + e_i + e_j)H^T \\ &= xH^T + e_iH^T + e_jH^T \\ &= 0 + h_i + h_j. \end{aligned}$$

Luego el síndrome de $y$ es diferente del vector cero ($s_y \neq 0$), porque $h_i$ y $h_j$ son diferentes de cero y linealmente independientes.

Si se produce sólo un error, entonces $y = x + e_i$, al hacer la misma operación se tiene la igualdad $yH^T = h_i$. El resultado de esta operación nos informa (en binario) de la posición donde ha tenido lugar el error y podemos corregirlo.

Esté código tiene la característica de poder corregir $t$ errores a todas las palabras del espacio que llegan al receptor, porque cada vector de longitud $n$ está en una de las esferas con centro en palabras del código. Por esta razón, Hamming es perfecto.

**Teorema 11.** HAM$(r)$ *es un código perfecto, es decir alcanza la cota Hamming.*

**Demostración**.

HAM$(r)$ tiene peso mínimo tres, es decir, $d_c = 3$, así que $t = 1$. Tenemos que $n = 2^r - 1$, $M = 2^{n-r}$. Por tanto

$$2^{n-r} \sum_{i=0}^{t=1} \binom{n}{i} = 2^{n-r}(1 + n) = 2^{n-r}(1 + 2^r - 1) = 2^{n-r+r} = 2^n. \ \nabla$$

Debido a que el código Hamming corrige sólo un error no es muy eficiente, por esta razón se recurre a los códigos extendidos los cuales permiten detectar y corregir más errores.

## 2.2. Código Hamming extendido

Si a un código Hamming se le añade un bit de paridad se aumenta la longitud y la distancia mínima. Este nuevo código se llama código Hamming extendido y se denota por EHAM$(r)$.

**Teorema 12.** *Una matriz de paridad para el código* EHAM$(r)$ *se denota* $H_E$ *de tamaño* $(r+1) \times (2^r)$ *y tiene la forma*

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ & & & 0 \\ & H & & \vdots \\ & & & 0 \end{pmatrix}$$

**Demostración**.

Se debe mostrar que $H_E$ tiene sus columnas linealmente independientes dos a dos. Como la matriz de paridad es la matriz generadora del código dual, se cumple que sus filas son linealmente independientes, luego al aumentar una columna de ceros a $H$ esta nueva matriz de tamaño $r \times (2^r)$ también tiene sus filas linealmente independientes. Ahora, al aumentar a esta matriz una fila de unos se tiene una matriz de tamaño $(r+1) \times (2^r)$, esta nueva fila no se puede obtener como combinación lineal de las anteriores, porque su última componente igual a uno no es igual a una suma de ceros. Por tanto, la matriz $H_E$ tiene sus filas linealmente independientes.

Sus columnas son linealmente independientes por pares, puesto que las columnas de $H$ lo son y como la última columna representa en binario el número 1, esta columna nunca será un múltiplo escalar de otros vectores.

Para $\hat{x} \in F_2^{n+1}$ se cumple que $\hat{x}H_E^T = \boldsymbol{0}$. En efecto, si $x \in$ HAM$(r)$ entonces $xH^T = \boldsymbol{0}$. El producto punto entre el vector $\hat{x} = (x_1, x_2, \ldots, x_n, x_{n+1})$ con la primera columna de $H_E^T$ es $(x_1 + x_2 + \ldots + x_n + x_{n+1})$ donde $x_{n+1} = \sum_{i=1}^{n} x_i$mód2, luego $[(x_1 + x_2 + \ldots + x_n) + \sum_{i=1}^{n} x_i]$mód2 $= 0$. Por tanto $\hat{x}H_E^T = \boldsymbol{0}$, entonces $H_E$ es una matriz de chequeo de paridad para el código Hamming extendido. $\nabla$

Al realizar la construcción de códigos Hamming extendidos se debe tener cuidado al aumentar los bits de paridad, para no reducir la distancia mínima debajo de 3.
A continuación se mencionan las propiedades más importantes del código Hamming. Considérese un código del tipo $[n, k, d_c]$ entonces,

1. La longitud es $n = 2^r - 1$, para $r = 2, 3, ...$
2. La dimensión $k$ del código cumple que $k = 2^r - 1 - r$.
3. El número de bits de chequeos de paridad es $r$.
4. La distancia mínima es $d_c = 3$.

## 2.3. Código cíclico

Los códigos cíclicos forman una subfamilia de códigos líneales o de bloques. En realidad, muchos de los códigos de bloques importantes descubiertos hasta la fecha son cíclicos o están bastante relacionados con este tipo de códigos. Una ventaja de los códigos cíclicos sobre la mayor parte de los otros tipos de códigos, es que son más fáciles de codificar, además poseen una estructura matemática perfectamente definida, la cual ha conducido al desarrollo de esquemas de decodificación muy eficientes para ellos. Se afirma que un código lineal será cíclico si cualquier corrimiento cíclico de una palabra del código produce de nuevo una palabra del código.

**Observación 11.** *De ahora en adelante se hace el convenio, que toda palabra $x$ en un código se representa de la siguiente forma $(x_0, x_1, ..., x_{n-1})$.*

**Definición 27.** *Un código lineal $C \subseteq F_2^n$ se llama cíclico si para todo $x = (x_0, x_1, ..., x_{n-1}) \in C$ su permutación cíclica $\sigma(x) = (x_{n-1}, x_0, ..., x_{n-2}) \in C$.*

En la definición hemos rotado todos los elementos del vector $x$ una unidad hacia la derecha. Por aplicación repetida de esta propiedad, vemos que el código es invariante por la rotación de $r$ elementos a la derecha, porque siempre se obtiene una palabra en el código $C$. Como una rotación de $n-1$ elementos hacia la derecha es igual que una rotación de un elemento

a la izquierda, el código también es invariante por rotaciones a la izquierda. En definitiva, un código cíclico es invariante por cualquier tipo de permutación cíclica de sus elementos.

**Teorema 13.** *Sea $x \in {F_2}^n$ un vector no nulo. El subespacio generado por los $n$ vectores $\sigma^i(x)$, con $i = 0, 1, \ldots, (n-1)$, obtenidos por rotaciones de $x$ es un código cíclico.*

**Demostración.**
Sea $C = gen\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^{n-1}(x)\}$. Probemos que $C$ es un código lineal.
$C$ es subespacio de ${F_2}^n$, además, si $z, w \in C$ por la forma en que se define $C$ existen escalares $a_0, \ldots, a_{n-1}$ y $b_0, \ldots, b_{n-1}$ en $F_2$ tales que

$$z = a_0\sigma^0(x) + \cdots + a_{n-1}\sigma^{n-1}(x) \text{ y } w = b_0\sigma^0(x) + \cdots + b_{n-1}\sigma^{n-1}(x).$$

Al sumar $z$ y $w$ se tiene

$$\begin{aligned} z + w =& a_0\sigma^0(x) + \cdots + a_{n-1}\sigma^{n-1}(x) + b_0\sigma^0(x) + \cdots + b_{n-1}\sigma^{n-1}(x) \\ =& (a_0 + b_0)\sigma^0(x) + \cdots + (a_{n-1} + b_{n-1})\sigma^{n-1}(x). \end{aligned}$$

Por tanto, $(z + w) \in C$, así $C$ es un código lineal.
Como los vectores $\sigma^i(x)$ generan el subespacio, un subconjunto de ellos forman una base. Sea $\left\{\sigma^0(x), \sigma^1(x), \ldots, \sigma^k(x)\right\}$ los elementos de la base, indizados en algún orden. Es claro que toda rotación de cualquier elemento de esta base es nuevamente un elemento de la forma $\sigma^i(x)$ que está contenido en el subespacio. $\nabla$

A continuación se describen los códigos cíclicos por medio de una estructura algebraica diferente. Se muestra que un código cíclico de longitud $n$ sobre $F_2$ consiste de todos los múltiplos de un polinomio generador $p(z)$, el cual es un polinomio mónico, irreducible de grado mínimo en el ideal y es un divisor de $z^n - 1$.
Se estudia los factores de $z^n - 1$, los cuales serán los generadores de los códigos cíclicos, además se asume que $n$ y 2 son primos relativos, así los ceros de $z^n - 1$ pertenecen a algún campo de extensión $F_{2^m}$, donde $m$ es el entero positivo más pequeño tal que $n$ divide a $2^m - 1$.

**Teorema 14.** *Si $n$ y 2 son primos entre sí, el polinomio $(z^n - 1) \in F_2[z]$ no posee raíces múltiples en.*

**Demostración.**

Sea $q(z) = (z^{n-1} + \ldots + z + 1)$. Es claro que $(z^n - 1) = (z - 1)q(z)$.

Las posibles raíces para $t(z) = z^n - 1$ son $z = 0$ y $z = 1$. Como $t(0) = 1$ entonces cero no es raíz, sin embargo $z = 1$ si es raíz. Mostremos que $z = 1$ no es raíz múltiple. Como $n$ es impar entonces $q(1) = 1\text{mód}2$. Esto muestra que 1 no es raíz de $q(z)$ y así $(z^n - 1)$ no posee raíces múltiples. $\nabla$

De ahora en adelante sólo se consideran códigos cíclicos de longitud $n$, donde se cumpla que el $\text{mcd}(2, n) = 1$. De esta forma, el grado del polinomio $z^n - 1$ y la característica del cuerpo a la que pertenecen sus coeficientes son primos entre sí y podemos aplicar el teorema anterior. La herramienta más importante en la descripción de los códigos cíclicos es el isomorfismo existente entre $F_2^n$ y un grupo de polinomios de grado a lo más $n - 1$ en $F_2[z]$, dado de la siguiente manera

**Teorema 15.** *Sea $\lambda \in F_2$, $c = (c_0, c_1, \ldots, c_{n-1}) \in F_2^n$. Existe un isomorfismo entre los espacios vectoriales $F_2^n$ y $R_n = F_2[z]/\langle z^n - 1\rangle$, definido así*

$$\begin{array}{|ll|}\hline \varphi : & F_2^n \to R_n \\ & c \mapsto \varphi(c) := c(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1} \\ \hline \end{array}$$

**Demostración.**

Probemos que $\varphi$ es función. Sean $b, c \in F_2^n$, donde $(b_0, b_1, \ldots, b_{n-1}) = (c_0, c_1, \ldots, c_{n-1})$ luego por igualdad de vectores $c_i = b_i$ para todo $0 \leq i \leq n - 1$, por tanto $\varphi(c) = \varphi(b)$.

Demostremos que $\varphi$ es inyectiva. Sean $c, b \in F_2^n$, supóngase que $\varphi(c) = \varphi(b)$ es decir, $c(z) = b(z)$. Por igualdad de polinomios los coeficientes son iguales, esto es, $c_i = b_i$ para todo $0 \leq i \leq n - 1$. Por tanto $(b_0, b_1, \ldots, b_{n-1}) = (c_0, c_1, \ldots, c_{n-1})$.

$\varphi$ es función sobreyectiva. En efecto, para todo polinomio $c(z) \in F_2[z]/\langle z^n - 1\rangle$, donde $c(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$, existe $c = (c_0, c_1, \ldots, c_{n-1})$ tal que $\varphi(c) = c(z)$.

Verifiquemos que $\varphi$ es un isomorfismo de espacios vectoriales.

1. $\varphi$ es lineal En efecto,

$$\begin{aligned}\varphi(b+c) &= \varphi((b_0+c_0,\ldots,b_{n-1}+c_{n-1}))\\ &= (b_0+c_0)+(b_1+c_1)z+\cdots+(b_{n-1}+c_{n-1})z^{n-1}\\ &= b_0+b_1z+\cdots+b_{n-1}z^{n-1}+c_0+c_1z+\cdots+c_{n-1}z^{n-1}\\ &= \varphi(b)+\varphi(c).\end{aligned}$$

2. $\varphi$ cumple la multiplicación por escalar.

$$\begin{aligned}\varphi(\lambda b) &= \varphi((\lambda b_0,\lambda b_1,\ldots,\lambda b_{n-1}))\\ &= \lambda b_0+\lambda b_1z+\cdots+\lambda b_{n-1}z^{n-1}\\ &= \lambda(b_0+b_1z+\cdots+b_{n-1}z^{n-1})\\ &= \lambda\varphi(b).\end{aligned}$$

Por lo tanto $\varphi$ es un isomorfismo de espacios vectoriales. $\nabla$
Ahora bien, consideremos el anillo cociente $R_n = F_2[z]/(z^n-1)$. Un elemento perteneciente a una clase de $R_n$ es un representante de ella, pero utilizando el algoritmo de la división observamos que hay un único representante de grado menor que $n$. En efecto, dado $q(z)$ un elemento de grado arbitrario, efectuamos la división euclídea, así

$$q(z) = c(z)(z^n-1)+r(z) \Rightarrow q(z)-r(z) = c(z)(z^n-1),$$

donde el polinomio $q(z) \cong r(z)\text{mód}(z^n-1)$ es de grado menor que $n$ y es único. Es importante tener presente que $(z^n-1)$ pertenece a la clase del cero por la forma en que está constituido $R_n$. Igualmente, todos los elementos de $F_2[z]/(z^n-1)$ pueden ser representados por polinomios de grado menor que $n$, es decir, las clases residuales del anillo cociente $F_2[z]/(z^n-1)$ tienen al siguiente conjunto de polinomios

$$\left\{x_0+x_1z+...+x_{n-1}z^{n-1} : x_i \in F_2, 0 \leq i \leq n-1\right\},$$

como un sistema de representaciones. Si $q(z) = x_0+x_1z+\ldots+x_{n-1}z^{n-1}$ es un elemento de $R_n$, lo podemos entender como el vector $x = (x_0,x_1,\ldots,x_{n-1})$, además, se introduce

la multiplicación de polinomios módulo $(z^n - 1)$ en la forma usual, es decir, si $q(z)$ es el representante de una clase lateral, $q(z) \in F_2[z]/(z^n - 1)$ y $g(z), h(z) \in F_2[z]$, entonces $g(z)h(z) = q(z)$, significa $g(z)h(z) \cong q(z)\text{mód}(z^n - 1)$.

Para desarrollar las propiedades algebraicas de los códigos cíclicos, es necesario asociarle a una palabra $x \in C \subseteq F_2^n$ su polinomio código $q(z) = x_0 + x_1 z + ... + x_{n-1}z^{n-1}$, siendo $q(z) \in S \subset F_2[z]/(z^n - 1)$. Para hacer un corrimiento hacia la derecha en un vector código, se realiza el producto entre $z$ y su polinomio correspondiente $q(z)$, donde la potencia de $z$ representa el número de desplazamientos en la palabra del código.

**Teorema 16.** *Si $q(z)$ es un polinomio código, entonces el polinomio*

$$q^i(z) = z^i q(z) mód(z^n + 1),$$

*también es un polinomio código para todo corrimiento cíclico $i$.*

**Demostración.**

Sea $C \subseteq F_2^n$ un código lineal y $(x_0, x_1, ..., x_{n-1})$ una palabra del código, a la cual le corresponde el polinomio $q(z) = x_0 + x_1 z + ... + x_{n-1}z^{n-1} \in S$. Al multiplicar $q(z)$ por $z^i$ se tiene

$$\begin{aligned} z^i q(z) =& z^i(x_0 + x_1 z + ... + x_{n-i}z^{n-i} + x_{n-i+1}z^{n-i+1} + ... + x_{n-1}z^{n-1}) \\ =& x_0 z^i + x_1 z^{i+1} + ... + x_{n-i}z^n + x_{n-i+1}z^{n+1} + ... + x_{n-1}z^{n+i-1} \\ =& x_{n-i}z^n + ... + x_{n-1}z^{n+i-1} + x_0 z^i + x_1 z^{i+1} + ... + x_{n-i-1}z^{n-1}. \end{aligned}$$

Ahora, como los $x_i \in F_2$, sumando y restando los términos $x_{n-i} + \cdots + x_{n-1}z^{i-1}$, se tiene

$$\begin{aligned} z^i q(z) = & x_{n-i}z^n + ... + x_{n-1}z^{n+i-1} + x_0 z^i + ... + x_{n-i-1}z^{n-1} + x_{n-i} \\ & + \cdots + x_{n-1}z^{i-1} + x_{n-i} + \cdots + x_{n-1}z^{i-1}. \end{aligned}$$

Manipulando los términos se obtiene

$$\begin{aligned} z^i q(z) =& x_{n-i} + \cdots + x_{n-1}z^{i-1} + x_0 z^i + ... + x_{n-i-1}z^{n-1} + x_{n-i}z^n \\ & + x_{n-i} + ... + x_{n-1}z^{n+i-1} + x_{n-1}z^{i-1} \end{aligned}$$

Sean los siguientes polinomios

$q^i(z) = x_{n-i} + ... + x_{n-1}z^{i-1} + x_0 z^i + x_1 z^{i+1} + ... + x_{n-i-}z^{n-1}$.

$h(z) = x_{n-i} + x_{n-i+1}z + ... + x_{n-1}z^{i-1}$.

Así, podemos expresar

$z^i q(z) = h(z)(z^n + 1) + q^i(z)$. Esto es, $q^i(z) = z^i q(z) \text{mód}(z^n + 1)$, donde $q^i(z) \in S$ corresponde a la palabra $(x_{n-i+1}, \ldots, x_{n-1}, x_0, x_1, \ldots, x_{n-i}) \in C$, que se obtiene al aplicar $i$ desplazamientos cíclicos a la palabra del código original. $\nabla$

Utilizando el isomorfismo establecido anteriormente, se muestra que los códigos cíclicos en $F_2^n$ corresponden a los ideales de $R_n$.

**Teorema 17.** *$C \subseteq F_2^n$ es un código cíclico si y sólo si $\varphi(C) = S$ es un ideal de $R_n$.*

**Demostración.**

$\Rightarrow$) Sea $C$ un código cíclico, $x \in C$ y $q(z) \in S$ su polinomio código. $S$ es cerrado bajo la suma porque su preimágen es un código cíclico. Probemos que para cualquier $a(z) \in R_n$ y $q(z) \in S$ se tiene que $a(z)q(z) \in S$.

$a(z)q(z) = (a_0 + a_1 z + a_2 z^2 + \ldots + a_{n-1}z^{n-1})q(z) = a_0 q(z) + a_1 z q(z) + \ldots + a_{n-1}z^{n-1}q(z)$.

Por el teorema 16, $a_i z^i q(z) \in S$ por ser un polinomio código, para cada $0 \leq i \leq (n-1)$. Como $S$ es cerrado bajo la suma $a(z)q(z) \in S$. Con lo cual se concluye que $S$ es un ideal.

$\Leftarrow$) Sea $S$ un ideal en $F_2[z]/(z^n - 1)$, $q(z) = x_0 + x_1 z + \cdots + x_{n-1}z^{n-1}$ el polinomio código correspondiente a la palabra $(x_0, x_1, \ldots, x_{n-1}) \in C$, por teorema 16, $zq(z)$ es un polinomio código, al cual le corresponde la palabra $(x_{n-1}, x_0, x_1, \cdots, x_{n-2}) \in C$. Por tanto $C$ es cíclico.$\nabla$

Veamos que $R_n = F_2[z]/(z^n - 1)$ es un anillo de ideales principales y que todo ideal tiene un generador particular.

**Teorema 18.** *Sea $C \subseteq F_2^n$ un código cíclico y $S$ el correspondiente ideal en $R_n$. Si $p(z)$ es el polinomio mónico de menor grado en $S$, entonces $p(z)$ está determinado en forma única, es un divisor de $z^n - 1$ y $S = \langle p(z) \rangle$.*

**Demostración.**

Sea $p(z)$ el polinomio mónico de menor grado en $S$ y $a(z)$ otro polinomio en $S$. Por el algoritmo de la división, existen $b(z)$ y $r(z)$ en $F_2[z]$ los cuales satisfacen

$$a(z) = p(z)b(z) + r(z),$$

donde el grado de $r(z)$ es menor que el grado de $p(z)$ o $r(z) = 0$. Luego

$$r(z) = a(z) - p(z)b(z),$$

y como $S$ es ideal, $r(z) \in S$, lo cual contradice la elección de $p(z)$, a menos que $r(z) = 0$, por lo cual $a(z) = p(z)b(z)$ y $p(z)$ genera el ideal.

Mostremos que $p(z)$ es único. Supóngase que existen $p(z)$ y $h(z)$ polinomios mónicos de grado mínimo en $S$. Así $p(z) - h(z) = q(z) \in S$, es un polinomio de menor grado que $p(z)$ y $h(z)$, lo cual es una contradicción. Por tanto existe un único polinomio mónico $p(z)$ de grado mínimo tal que $S = \langle p(z) \rangle$.

Probemos que $p(z)$ es un divisor de $(z^n - 1)$. Al dividir $(z^n - 1)$ entre el generador $p(z)$ se obtiene

$$(z^n - 1) = c(z)p(z) + r(z),$$

luego $0 \cong p(z)c(z) + r(z)$ donde el grado $r(z)$ es menor que el grado de $p(z)$. Luego $-p(z)c(z) = r(z) \in S$, esto es una contradicción a menos que $r(z)$ sea identicamente cero. Por tanto $p(z)$ divide a $(z^n - 1)$. $\nabla$

**Observación** En adelante se usará la terminología de vectores $(x_0, x_1, ..., x_{n-1}) \in C$ y polinomios $q(z) = \sum_{i=0}^{n-1} x_i z^i$, $q(z) \in S$ sobre $F_2$ similarmente. Podemos interpretar al código cíclico $C$ como el subconjunto $S$ del anillo factor $F_2[z]/(z^n - 1)$.

Sea $(z^n - 1) \in F_2[z]$, $(z^n - 1) = q_1(z)q_2(z)...q_t(z)$ la descomposición del polinomio $(z^n - 1)$ en factores irreducibles diferentes. Podemos encontrar todos los códigos cíclicos de longitud $n$, escogiendo (entre todas las posibles formas) uno de los $2^t$ factores de $(z^n - 1)$ como polinomio generador $p(z)$ y definiendo el correspondiente código al conjunto formado por todos los múltiplos de $p(z)$mód$(z^n - 1)$.

**Definición 28.** *El código cíclico generado por el polinomio irreducible $q_i(z)$ es llamado un* ***código cíclico maximal****.*

**Definición 29.** *El código cíclico generado por el polinomio*

$$\begin{aligned}\hat{q}_i(z) =& q_1(z)\cdots q_{i-1}(z)q_{i+1}(z)\cdots q_t(z)\\ \hat{q}_i(z) =& \frac{(z^n-1)}{q_i(z)},\end{aligned}$$

*se llama* ***código cíclico minimal****.*

**Ejemplo 2.**

Sea $V = F_2^3$ y $R_3 = F_2[z]/(z^3-1)$. Localicemos simultáneamente los códigos cíclicos en $V$ e ideales en $R_3$. La factorización de $(z^3-1) = (1+z)(1+z+z^2)$, da como resultado $2^2$ códigos cíclicos. La siguiente tabla muestra tales códigos

| Código | Vectores | Ideal |
|---|---|---|
| $C_1$ | $(0,0,0)$ | $\langle 0\rangle$ |
| $C_2$ | $(0,0,0)$ | $\langle 1+z+z^2\rangle$ |
| | $(1,1,1)$ | |
| $C_3$ | $(0,0,0)$ | $\langle 1+z\rangle$ |
| | $(1,1,0)$ | |
| | $(0,1,1)$ | |
| | $(1,0,1)$ | |
| $C_4$ | V | $R_3$ |

### 2.3.1. Codificación

Considérese los polinomios

$$\begin{aligned}m(z) =& m_0 + m_1 z, ..., m_{k-1}z^{k-1} \in F_2^k,\\ b(z) =& b_0 + b_1 z, ..., b_{n-k-1}z^{n-k-1} \in F_2^n,\end{aligned}$$

donde $m(z)$ es el polinomio del vector mensaje y $b(z)$ es el polinomio que tiene como coeficientes los bits de paridad.

Para codificar la secuencia del mensaje $(m_0, m_1, ..., m_{k-1})$ en un código cíclico sistemático $[n, k, d_c]$, donde los bits de información se transmiten de forma inalterada, se encuentra el

polinomio código $q(z)$ el cual es la suma del polinomio de paridad y el polinomio mensaje con $n-k$ corrimientos, es decir, $q(z)$ se expresa $q(z) = b(z) + z^{n-k}m(z)$ y se encuentra con el siguiente algoritmo

- Sea $m(z)$ el polinomio mensaje, al hacer $n-k$ corrimientos se obtiene el polinomio $z^{n-k}m(z)$.
- Se divide $z^{n-k}m(z)$ entre el polinomio generador $p(z)$, luego

  $$z^{n-k}m(z) = p(z)a(z) + b(z),$$

  donde $b(z)$ es el residuo, llamado polinomio de paridad.
- Al sumar módulo 2 se obtiene $p(z)a(z) = z^{n-k}m(z) + b(z)$.
- Como el código es generado por $p(z)$, se cumple que $q(z) = p(z)a(z)$, para algún $a(z) \in F_2[z]$.
- Reemplazando $q(z) = b(z) + z^{n-k}m(z)$.
- Dado $q(z)$ se encuentra su vector de $n$ componentes correspondiente.

**Ejemplo 3.**

Sea $M \subset F_2^4$ el conjunto formado por las palabras mensaje. Sea $S \subset F_2[z]/\langle z^7-1\rangle$. Pero $(z^7-1) = (z-1)(1+z+z^3)(1+z^2+z^3)$ donde $S = \langle 1+z+z^3\rangle$. Hallar la palabra código correspondiente al vector mensaje $m = (1,0,1,1) \in M$.

- El polinomio mensaje es $m(z) = 1 + z^2 + z^3$ y el producto

  $z^{7-4}m(z) = z^3(1+z^2+z^3) = z^3 + z^5 + z^6$.
- Al dividir $z^3m(z)$ entre $p(z)$ se tiene

$$\begin{aligned} z^6 + z^5 + z^3 =& (z^3+z+1)(z^3+z^2+z+1)+1 \\ =& z^6 + z^5 + z^3 + 1. \end{aligned}$$

- El polinomio código es $q(z) = 1 + z^3 + z^5 + z^6$
- La palabra código es $q = (1, 0, 0, 1, 0, 1, 1)$.

Las primeras tres coordenadas corresponden al vector de paridad y las cuatro últimas forman el vector mensaje $m$.

### 2.3.2. Matriz generadora del código

Los elementos de un código cíclico $[n, k, d_c]$ pueden obtenerse multiplicando cada polinomio mensaje de grado $k-1$ por un polinomio fijo $p(z)$ de grado $n-k$. Si $p(z)$ es el polinomio generador de un código cíclico $S$ de longitud $n$ y de grado $n-k$, entonces el conjunto de polinomios $\{p(z), zp(z), \ldots, z^{k-1}p(z)\}$ corresponden a palabras del código y forman una base para $S$, además la dimensión de $S = \langle p(z) \rangle$ es $k$. El siguiente teorema formaliza estas afirmaciones.

**Teorema 19.** *Si $p(z) = p_0 + p_1 z + \ldots + p_{n-k} z^{n-k}$ es un divisor de $z^n - 1$ y $S$ es un código cíclico generado por $p(z)$, entonces $S$ tiene dimensión $k$ y*

$$G = \begin{pmatrix} p_0 & p_1 & \ldots & p_{n-k} & & 0 & 0 & \ldots & 0 \\ 0 & p_0 & p_1 & \ldots & & p_{n-k} & 0 & \ldots & 0 \\ \vdots & \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \ldots & 0 & p_0 & p_1 & \ldots & p_{n-k} \end{pmatrix}$$

es una matriz generadora para $S$.

**Demostración.**

Se debe garantizar que las filas de $G$ forman una base para el código.

Las $k$ filas de $G$, $p(z), zp(z), \ldots, z^{k-1}p(z)$ corresponden a permutaciones de $p(z)$. Supongamos que estos polinomios son linealmente dependientes, luego existen escalares $a_i \in F_2$, $0 \leq i \leq k-1$, no todos cero tales que

$$a_0 p(z) + a_1 z p(z) + \ldots + a_{k-1} p(z) z^{k-1} = (a_0 + a_1 z + \ldots + a_{k-1} z^{k-1}) p(z) = 0$$

Pero este producto tiene grado menor que $n$, así no puede ser cero módulo $(z^n - 1)$, a menos que cada $a_i = 0$, lo que contradice la elección de los $a_i$. Por tanto el conjunto de polinomios es linealmente independiente.

Ahora se prueba que las filas de $G$ generan a $S$. Sea $s(z) \in S$, el cual puede expresarse como $s(z) = c(z)p(z)$ donde el grado de $c(z)$ es menor o igual que $(k-1)$. Luego,

$$s(z) = c(z)p(z) = (c_0 + c_1 z + \ldots + c_{k-1}z^{k-1})p(z)$$
$$s(z) = c_0 p(z) + c_1 z p(z) + \ldots + c_{k-1}z^{k-1}p(z).$$

Esto muestra que $s(z)$ se puede escribir como combinación lineal de las filas de $G$. Así, las filas de la matriz $G$ forman una base para $S$. Por tanto $S$ tiene dimensión $k$.$\nabla$

Por lo anterior, para codificar una secuencia de información $(m_0, m_1, \ldots, m_{k-1})$ se hace el producto $s = mG$ o también

$$s(z) = (m_0 + m_1 z + \cdots + m_{k-1}z^{k-1})p(z).$$

Al factorizar $z^n - 1$ en $F_2[z]$ se sabe cuantos códigos cíclicos de longitud $n$ hay sobre $F_2$. Además, es posible saber la dimensión del código cíclico con el grado del factor (dado que este polinomio factor es el generador del código), también dicho factor da lugar a una matriz generadora explícita.

### 2.3.3. Código Dual

Sea $p(z)$ el polinomio generador de un código cíclico $S$. Si $p(z)$ tiene grado $n-k$, divide a $z^n - 1$, así que $(z^n - 1) = p(z)h(z)$, entonces $S$ tiene dimensión $k$ y $h(z)$ tiene grado $k$. Como $h(z)$ es un factor de $z^n - 1$, entonces es el generador de un código cíclico de dimensión $n-k$. Como $S^{\perp}$ es el código dual de $S$, su dimensión es $n-k$ y en verdad sería conveniente si $h(z)$ fuera su polinomio generador, pero no hay una razón para que esto sea así, lo sorprendente es que el polinomio generador del código dual puede ser descrito en términos de $h(z)$. A este polinomio se le denomina polinomio de paridad para $S$.

**Lema 2.** *Sea $S = \langle p(z) \rangle$ un código cíclico y $h(z)$ el polinomio de paridad para $S$. Un polinomio $q(z) \in S$ si y sólo si $q(z)h(z) \cong 0 mód (z^n - 1)$.*

**Demostración.**

$\Rightarrow$) Si $c(z) \in S$ entonces $c(z) \cong p(z)a(z)\text{mód}(z^n - 1)$. Multiplicando por $h(z)$ se tiene

$$c(z)h(z) \cong p(z)a(z)h(z)\text{mód}(z^n - 1).$$

Como $p(z)h(z)\text{mód}(z^n - 1) \cong 0\text{mód}(z^n - 1)$ entonces $c(z)h(z) \cong 0\text{mód}(z^n - 1)$ $\nabla$

$\Leftarrow$) Supóngase que $q(z)h(z) \cong 0\text{mód}(z^n - 1)$. Al dividir $q(z)$ entre $p(z)$ se tiene

$$q(z) = p(z)t(z) + r(z),$$

donde el grado de $r(z)$ es menor que el de $p(z)$. Multiplicando por $h(z)$ la igualdad

$$q(z)h(z) = (t(z)p(z) + r(z))h(z)$$
$$0 mód(z^n - 1) \cong t(z)p(z)h(z) + r(z)h(z)$$

Como $p(z)h(z) \cong 0\text{mód}(z^n - 1)$ entonces $r(z)h(z) \cong 0\text{mód}(z^n - 1)$. Pero el grado de $r(z)h(z)$ es menor que $n$ y así necesariamente $r(z) = 0$ y $q(z)$ es múltiplo del generador. Por lo tanto $q(z) \in S.\nabla$

Los resultados anteriores muestran la importancia del polinomio de chequeo de paridad, el cual permite definir el código cíclico de otra forma, además por medio de este, es posible obtener el generador del código dual.

**Definición 30.** *Sea $c(z) \in F_2[z]$, donde $c(z) = c_0 + c_1 z + \ldots + c_r z^r$ y $r = n - k$. Se llama polinomio recíproco de $c(z)$, al polinomio $\bar{c}(z) = z^r c(z^{-1})$, es decir,*

$$\bar{c}(z) = \sum_{i=0}^{r} c_{r-i} z^i = c_r + c_{r-1} + \ldots + c_0 z^r.$$

**Teorema 20.** *Si $p(z)h(z) = (z^n - 1)$ en $F_2[z]$ y $p(z)$ es el polinomio generador del código $S$, entonces el polinomio $\bar{h}(z) = h_k + h_{k-1}z + h_{k-2}z^2 + \cdots + h_0 z^k$ es el generador del código cíclico $S^{\perp}$. Además si $h(z) = h_0 + h_1 z + \ldots + h_k z^k$, entonces una matriz de chequeo de paridad de tamaño $(n - k) \times n$ para $S$ es*

$$H = \begin{pmatrix} h_k & h_{k-1} & \ldots & h_0 & 0 & \ldots & 0 \\ 0 & h_k & \ldots & h_1 & h_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \ldots & h_k & h_{k-1} & \ldots & h_1 & h_0 \end{pmatrix}$$

**Demostración.**

Para comprobar que $H$ es una matriz de chequeo de paridad es suficiente mostrar que para todo $c(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$, con $c(z) \in S$, su respectivo vector código $c$ cumple que $cH^T = 0$. Sea $c(z) \in S$, luego $c(z)h(z) \cong 0 mód(z^n - 1)$, donde $h(z)$ es el polinomio de paridad de $S$ Como

$$c(z)h(z) = q(z)(z^n - 1), \tag{2.1}$$

el producto $c(z)h(z)$ tiene grado a lo más $(n - 1 + k)$, luego $q(z)$ tiene grado a lo más $k - 1$. Por la igualdad 2,1, $c(z)h(z)$ se tiene que los coeficientes de $z^k, \ldots, z^{n-1}$ son cero, donde estos se expresan mediante las respectivas combinaciones lineales

$$\begin{aligned} c_0 h_k + c_1 h_{k-1} + \cdots + c_k h_0 &= 0 \\ c_1 h_k + c_2 h_{k-1} + \cdots + c_{k+1} h_0 &= 0 \\ &\vdots \\ c_{n-k-1} h_k + c_{n-k} h_{k-1} + \cdots + c_{n-1} h_0 &= 0 \end{aligned}$$

Estos coeficientes son precisamente el producto escalar del vector $c$ por cada fila de la matriz $H$. De esta forma, $c = (c_0, c_1, \ldots, c_{n-1})$ es ortogonal a cada una de las $n - k$ columnas de la matriz $H^T$ y así $cH^T = 0$.

Ahora se garantiza que el polinomio recíproco de $h(z)$ es un factor de $(z^n - 1)$. Sea $(z^n - 1) = p(z)h(z)$. Como

$$(z^{-n} - 1) = p(z^{-1})h(z^{-1}) = \frac{1 - z^n}{z^n},$$

se dan las igualdades siguientes

$$\begin{aligned} 1 - z^n &= z^n p(z^{-1}) h(z^{-1}) \\ &= z^{n-k} p(z^{-1}) z^k h(z^{-1}) \\ &= \bar{p}(z)\bar{h}(z). \end{aligned}$$

Pero en $F_2[z]$ se cumple que $1 - z^n = z^{-1}$ luego, $\bar{h}(z)$ divide a $z^{-1}$. Por tanto $\bar{h}(z)$ genera un código cíclico.

Ahora se prueba que $\bar{h}(z)$ genera a $S^{\perp}$. Como $H$ es la matriz generadora del código $S^{\perp}$, para todo $q(z) \in S^{\perp}$ se satisface que $q = mH$ donde $q$ denota el vector código correspondiente al polinomio $q(z)$. Así, $q(z) = m(z)\bar{h}(z)$ y por tanto $\left\langle \bar{h}(z) \right\rangle = S^{\perp}$.$\nabla$

### 2.3.4. Codificación sistemática

Hasta aquí se ha explicado la estructura general de los códigos cíclicos. En esta sección se muestra un enfoque alternativo de algunos conceptos, dado que el método de codificación usado anteriormente produce codificaciones no sistemáticas. Analizaremos ahora otro método, basado exclusivamente en el álgebra del anillo de polinomios, que da lugar a codificaciones sistemáticas y puede simplificar alguno de los desarrollos.

**Teorema 21.** *la aplicación $\phi$ es una función de codificación para el código con polinomio generador $p(z)$, siendo $r = n - k$ su grado.*

$$\boxed{\begin{array}{ll} \phi : R_k & \rightarrow R_n \\ q(z) & \rightarrow \phi(q(z)) = z^r q(z) - (z^r q(z) \textit{módp}(z)) \end{array}}$$

**Demostración.** Notemos que $\phi(q(z))$ es el cocientye de dividir $z^r q(z)$ entre $p(z)$.

$\phi$ es función. Sean $r(z)$, $t(z)$ polinomios en $R_k$. Supóngase que $r(z) = t(z)$. Luego $z^r r(z) = z^r t(z)$, con lo que $z^r r(z) - (z^r r(z) \textit{modp}(z)) = z^r t(z) - (z^r t(z) \textit{modp}(z))$ y por tanto $\phi(r(z)) = \phi(t(z))$.

La imagen de la aplicación está contenida en el código. En efecto, al hacer la división de $z^r q(z)$ entre $p(z)$ se tiene $z^r q(z) = c(z)p(z) + w(z)$ con el grado de $w(z)$ menor que el grado de $p(z)$, donde $w(z) = z^r q(z)$mód$p(z)$. Es claro que $z^r q(z) - w(z) = c(z)p(z)$, por tanto, la imagen de la aplicación pertenece al $\langle p(z) \rangle$.

Ahora mostremos que $\phi$ es una aplicación lineal. $\phi$ es lineal. Sean $q(z) = \sum_{i=0}^{k-1}(q_i z^i)$ y

$t(z) = \sum_{i=0}^{k-1}(t_i z^i)$ en $R_k$. Luego

$$\begin{aligned}
\phi(q(z)+t(z)) &= \phi\left(\sum_{i=0}^{k-1}(q_i z^i + t_i z^i)\right) \\
&= z^r\left(\sum_{i=0}^{k-1}(q_i z^i + t_i z^i)\right) - \left[z^r\left(\sum_{i=0}^{k-1}(q_i z^i + t_i z^i)\right)\right] mód(p(z) \\
&= z^r(q(z)+t(z)) - (z^r(q(z)+t(z)))módp(z) \\
&= z^r q(z) + z^r t(z) - z^r q(z)módp(z) - z^r t(z)módp(z) \\
&= z^r q(z) - z^r q(z)módp(z) + z^r t(z) - z^r t(z)módp(z) \\
&= \phi(q(z)) + \phi(t(z)).
\end{aligned}$$

Probemos que $\phi$ es inyectiva. Para eso se muestra que el Kernel de $\phi$ está formado por el polinomio cero. Sea $q(z) \in R_k$ con $q(z) \neq 0$, ahora $\phi(q(z)) = z^r q(z) - (z^r q(z)\text{mód } p(z))$. Sea $w(z) = z^r q(z)\text{mód}p(z)$ con grado de $w(z)$ menor que $r$ y el grado de $z^r q(z)$ mayor que $r$ (y menor que $n$), luego la diferencia es no nula, es decir $\phi(q(z))$ es distinto de cero, por tanto $Ker(\phi) = \{0\}$ lo que implica que la función es inyectiva.
Sea $\alpha$ en $F_2$. $\phi(\alpha q(z)) = \alpha\phi(q(z))$. En efecto,

$$\begin{aligned}
\phi(1q(z)) &= \phi(q(z)) \\
&= 1\phi(q(z))
\end{aligned}$$

Además,

$$\begin{aligned}
\phi(0q(z)) &= 0 \\
&= 0\phi(q(z)).
\end{aligned}$$

Para códigos cíclicos sistemáticos es más útil trabajar con una matriz generadora de la forma $G = (A, I)$. Recordemos que las filas de $G$ son los corrimientos cíclicos del polinomio generador $p(z)$. Dichos corrimientos, son precisamente las filas de $A$, por lo cual se escribe $A_i = (z^{i+r} mod p(z))$ donde $A_i$ es la fila i-ésima de la matriz $A$.
La matriz generadora asociada a la anterior aplicación se llama matriz sistemática del código. Esta se escribe

$$G=\begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0r-1} & 1 & 0 & 0 & \cdots & 0 \\ p_{10} & p_{11} & \cdots & p_{1r-1} & 0 & 1 & 0 & \cdots & 0 \\ p_{20} & p_{21} & \cdots & p_{2r-1} & 0 & 0 & 1 & \cdots & 0 \\ & & \vdots & & & & & \vdots & \\ p_{m-10} & p_{m-21} & \cdots & p_{m-1r-1} & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

**Corolario 7.** *Dada la matriz generadora $G = (A, I)$ en forma sistemática, su matriz de paridad $H = (I, -A^T)$ de tamaño $r \times n$ donde $r = n - k$, cumple $h_i = z^i mod p(z)$, siendo $h_i$ la columna i-ésima. (la numeración empieza en cero).*

**Demostración.**

Si $0 \leq i < r$, entonces $h_i = z^i$mód $p(z) = z^i$ y se obtiene la matriz identidad. La columna $h_r$ se obtiene trasponiendo la fila 0 de la matriz $A$ y cambiando el signo, entonces $h_r = -(A_0)^T = z^r$mód$p(z)$. La siguiente columna se obtiene así

$$h_{r+1} = -(A_1)^T = z^{r+1}\text{mód}p(z).$$

Se sigue el mismo proceso para obtener las columnas posteriores. Luego para toda columna $h_i$ de $H$ se tiene $h_i = z^i$mód$p(z)$ con $0 \leq i \leq n-1$.

$$H = \begin{pmatrix} 1 & 0 & \cdots & 0 & p_{00} & p_{10} & \cdots & p_{m-10} \\ 0 & 1 & \cdots & 0 & p_{01} & p_{11} & \cdots & p_{m-20} \\ & & \vdots & & & & & \vdots \\ 0 & 0 & \cdots & 1 & p_{0r-1} & p_{1r-1} & \cdots & p_{m-1r-1}. \end{pmatrix}$$

**Ejemplo 4.**

Sea $S \subseteq F_2[z]/\langle z^7 - 1\rangle$ donde $S = \langle z^3 + z + 1\rangle$. Haciendo uso de la aplicación lineal $\phi : R_3 \to R_7$ se calcula la matriz generadora asociada a la aplicación lineal para este código cíclico, encontrando las imagenes de la base $B = \{1, z, z^2, z^3\}$ de $R_3$.

$$\begin{aligned} \phi(1) &= z^3 - (z^3 mód[z^3 + z + 1]) = 1 + z + z^3 \\ \phi(z) &= z^3 z - (z^4 mód[z^3 + z + 1]) = z + z^2 + z^4 \\ \phi(z^2) &= z^3 z^2 - (z^5 mód[z^3 + z + 1]) = 1 + z + z^2 + z^5 \\ \phi(z^3) &= z^3 z^3 - (z^6 mód[z^3 + z + 1]) = 1 + z^2 + z^6. \end{aligned}$$

Las filas de la matriz $G$ son los vectores correspondientes a los polinomios obtenidos. Por tanto la matriz generadora es

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

y la matriz de paridad $H$ se puede construir directamente aplicando $H = (I, -A^T)$. Esto es,

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

### 2.3.5. Cálculo del síndrome

Supóngase que la palabra de código $(c_0, c_1, \ldots, c_{n-1})$ se transmite por un canal ruidoso originando una palabra recibida $(y_o, y_1, \ldots, y_{n-1})$. Recordemos que el primer paso en la decodificación de un código de bloques lineales consiste en calcular el síndrome $s_y = yH^T$ de la palabra recibida $y$. Cuando el síndrome es cero, no hay errrores de transmisión en la palabra recibida. Pero, si el síndrome es distinto de cero, la palabra recibida contiene errores de transmisión que es necesario corregir. En el caso de un código cíclico en forma sistemática, el síndrome puede calcularse con facilidad, haciendo uso de la matriz de paridad $H$, pero debido a su estructura, el síndrome se puede calcular sin realizar productos de matrices, utilizando únicamente operaciones con polinomios.

**Teorema 22.** *Sea $S = \langle p(z) \rangle$. Si al transmitir $q(z) \in S$ se cometen errores, el síndrome $s_y(z)$ de la palabra recibida $y(z)$ es*

$$s_y(z) = y(z) m\acute{o}d(p(z)).$$

**Demostración.**

De la definición de síndrome dada en el capítulo uno, se tiene $s_y = yH^T$, donde el vector $y$ es la correspondiente n-upla del polinimio $y(z) = y_0 + y_1 z + \cdots + y_{n-1}z^{n-1}$. Sea $h_i$ la columna i-ésima de la matriz de paridad $H$. Así,

$$s_y = yH^T = (y_0 h_0, y_1 h_1, \ldots, y_{n-1} h_{n-1}).$$

Pero por la estructura de $H$, se tiene que $h_i = z^i \text{mód} p(z)$, por lo cual aplicando la linealidad de la operación módulo tenemos que su polinomio correspondiente viene dado por $s_y(z) = (y_0 + y_1 z + y_2 z^2 + \cdots + y_{n-1}z^{n-1}) \text{mód} p(z)$ que es el resultado que se deseaba mostrar. $\nabla$

**Teorema 23.** *Sea $S$ un código cíclico de distancia mínima $d_c \geq 2t + 1$. Si en una transmisión $g(z) \to q(z)$ se producen menos de $t$ errores y el síndrome tiene peso menor o igual que $t$, entonces el error cometido es exactamente igual que el síndrome.*

**Demostración.**

Denotemos por $e(z)$ el polinomio error y por $s_q(z)$ el síndrome de $q(z)$; se cumple que $e(z) = q(z) - g(z)$ y $q(z) = c(z)w(z) + s_q(z)$. Reemplazando $q(z)$ en la primera igualdad se tiene $e(z) = c(z)w(z) + s_q(z) - g(z)$ por lo tanto $e(z) - s_q(z) = (c(z)w(z) - g(z)) \in S$. Como $e(z) - s_q(z)$ pertenece a $S$, encontrando sus vectores correspondientes se tiene que $(e - s_q) \in C$, pero el peso de esta resta cumple $w(e - s_q) \leq w(e + s_q) \leq 2t < d_c$, lo cual es una contradicción dado que distancia mínima $d_c = w_c$, entonces no puede existir un vector en $C$ con esta condición por lo cual la resta es necesariamente nula $e - s_q = 0$, en conclusión $e(z) = s_q(z)$. $\nabla$

### 2.3.6. Generadores Idempotentes

Se seguirá utilizando la condición que $n$ y 2 son primos relativos, lo cual implica que $z^n - 1$ se descompone en factores distintos. Estos factores son los generadores de diferentes códigos cíclicos y su grado brinda información de la dimensión del código, además dichos ideales pueden tener varios generadores. Tal es el caso de los llamados generadores idempotentes los cuales se describen a continuación.

**Definición 31.** *Un elemento $\mu(z)$ de un anillo se llama idempotente si*

$$\mu(z)^2 = \mu(z) = \mu(z^2).$$

**Proposición 15.** *Sea $S = \langle p(z) \rangle$ un código cíclico sobre $F_2$. Existe un único elemento $\mu(z)$ idempotente que genera el ideal.*

**Demostración.**

Mostremos que existe un polinomio idempotente. Sea $p(z)$ el generador del ideal (polinomio mónico de grado mínimo). Considérese la descomposición $z^n - 1 = p(z)h(z)$. Como $n$ y la característica del campo son primos relativos, resulta que $p(z)$ y $h(z)$ también son primos. Al aplicar el lema de Euclides se tiene que

$$1 = a(z)p(z) + b(z)h(z),$$

además

$$p(z)h(z) \cong 0 mód(z^n - 1) \tag{2.2}$$

Denótese $\mu(z)$ al producto $a(z)p(z)$, si multiplicamos la identidad anterior por $\mu(z)$ se obtiene

$$\begin{aligned} \mu(z) &= \mu(z)\mu(z) + b(z)h(z)\mu(z) \\ &= \mu(z)\mu(z) + b(z)h(z)p(z)a(z), \end{aligned}$$

Por la ecuación (2,2) el último sumando es nulo, así que $\mu(z) = \mu(z)\mu(z)$. Por tanto $\mu(z)$ es un elemento idempotente.

Sea $c(z)$ un elemento cualquiera del código, donde $c(z) = p(z)t(z)$. Como $p(z)$ y $h(z)$ son primos relativos, entonces existen polinomios $a(z)$ y $b(z)$ tales que $1 = a(z)p(z)+b(z)h(z)$. Multiplicando la identidad anterior por $c(z)$ se tiene

$$\begin{aligned} c(z) &= a(z)p(z)c(z) + b(z)h(z)c(z) \\ &= \mu(z)c(z) + b(z)h(z)p(z)t(z), \end{aligned}$$

de donde $c(z) = \mu(z)c(z)$, por lo cual $\mu(z)$ es otro generador del ideal, además se comporta como el neutro.

El elemento $\mu(z)$ es único. En efecto, si otro elemento idempotente genera el mismo ideal, también se comportará como un elemento neutro para el producto, de donde se deduce la unicidad.$\nabla$

La prueba anterior muestra una forma de encontrar el generador idempotente $\mu(z)$ para un código cíclico $S$ a partir del polinomio generador $p(z)$, además es posible encontrar $p(z)$ si se conoce $\mu(z)$, como lo muestra el siguiente teorema.

**Teorema 24.** *Si $S$ es un código cíclico en $F_2[z]/(z^n-1)$ con polinomio generador idempotente $\mu(z)$, entonces el polinomio generador $p(z)$ de $S$ es igual al $mcd(\mu(z),(z^n-1)$.*

**Demostración.**

Probemos que $d(z) = mcd(\mu(z),(z^n-1))$ es un divisor de $p(z)$. Como $d(z)$ divide a $\mu(z)$ entonces existe $k(z)$ en $F_2[z]$ tal que $\mu(z) = d(z)k(z)$, además $S = \langle\mu(z)\rangle$. Luego, para todo $m(z) \in S$ se cumple que $m(z) = \mu(z)r(z)$. Reemplazando $\mu(z)$ se tiene que $m(z) = d(z)k(z)r(z)$, por tanto $m(z) \in \langle d(z)\rangle$. Como $S$ es un ideal entonces $S \subset \langle d(z)\rangle$ y $d(z)$ divide a $p(z)$.

$p(z)$ divide a $(z^n-1)$ y a $\mu(z)$, luego $p(z)$ divide a $d(z)$ por tanto $\langle d(z)\rangle \subseteq \langle p(z)\rangle$. Así $d(z) = p(z)$. $\nabla$

**Ejemplo 5.** *la siguiente tabla muestra todos los códigos cíclicos de longitud siete sobre $F_2$ junto con sus polinomios generadores $p_i(z)$ y sus generadores idempotentes $\mu_i(z)$.*

| i | dim | $p_i(z)$ | $\mu_i(z)$ |
|---|---|---|---|
| 0 | 0 | $p_0(z) = 1+z^7$ | $\mu_0(z) = 0$ |
| 1 | 1 | $p_1(z) = z(1+z+z^3)(1+z^2+z^3)$ | $\mu_1(z) = p_1(z) = 1+z+z^2+\cdots+z^6$ |
| 2 | 3 | $p_2(z) = (1+z)(1+z+z^3)$ | $\mu_2(z) = z^3p_2(z) = 1+z^3+z^5+z^6$ |
| 3 | 3 | $p_3(z) = (1+z)(1+z^2+z^3)$ | $\mu_3(z) = p_3(z) = 1+z+z^2+z^4$ |
| 4 | 4 | $p_4(z) = 1+z+z^3$ | $\mu_1(z)+\mu_2(z) = zp_5(z) = z+z^2+z^4$ |
| 5 | 4 | $p_5(z) = 1+z^2+z^3$ | $\mu_1(z)+\mu_3(z) = z^3p_6(z) = z^3+z5+z^6$ |
| 6 | 6 | $p_6(z) = 1+z$ | $\mu_2(z)+\mu_3(z) = p_4(z)(z+z^3+z^5)$ |
| 7 | 7 | $p_7(z) = 1$ | $\mu_7(z) = 1$ |

**Teorema 25.** *Sea $S$ un código cíclico con parámetros $[n, k, d_c]$ con generador idempotente $\mu(z) = \sum_{1=0}^{n-1} \mu_i z^i$. Entonces la siguiente matriz de tamaño $k \times n$, es una matriz generadora para $S$.*

$$G = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_{n-2} & \mu_{n-1} \\ \mu_{n-1} & \mu_0 & \mu_1 & \cdots & \mu_{n-3} & \mu_{n-2} \\ \vdots & & & \vdots & & \\ \mu_{n-k+1} & \mu_{n-k+2} & \mu_{n-k+3} & \cdots & \mu_{n-k-1} & \mu_{n-k} \end{pmatrix}$$

**Demostración.**

Mostrar que $G$ es una matriz generadora es equivalente a probar que el conjunto de filas $\{\mu(z), z\mu(z), \ldots, z^{k-1}\mu(z)\}$ es una base para $S$.

Considérese la siguiente combinación lineal igualada a cero

$$a_0\mu(z) + a_1 z\mu(z) + \cdots + a_{k-1}z^{k-1}\mu(z) = 0,$$

donde los $a_i \in F_2$, $0 \leq i \leq (k-1)$. Factorizando $\mu(z)$ se tiene la igualdad

$$(a_0 + a_1 z + \cdots + a_{k-1}z^{k-1})\mu(z) = 0.$$

Denotemos por $a(z) = a_0 + a_1 z + \cdots + a_{k-1}z^{k-1}$. Como el elemento idempotente actua como la identidad se tiene $\mu(z)a(z)\mu(z) = a(z) = 0$. Por tanto el conjunto es linealmente independiente.

Los corrimientos de $\mu(z)$ generan el código. En efecto, para todo $a(z) \in S$ se cumple que $a(z) = a(z)\mu(z)$. Por tanto, el conjunto es una base de $S$ y la matriz $G$ es una matriz generadora para el código $S$. $\nabla$

## 2.4. Distancia mínima de un código cíclico

En cualquier código es importante establecer la distancia mínima con el fin de determinar su capacidad para corregir y detectar errores. Para los códigos cíclicos hacemos uso de la cota BCH. Antes de introducirla retomamos algunos conceptos de la teoría de campos.

Si $f(z)$ es un polinomio en $F_2[z]$ y $\alpha$ es una raíz en algún campo de extensión $F_{2^t}$ entonces $\alpha^2$ es también raíz de $f(z)$ en $F_2[z]$, aplicando repetidamente el resultado se tiene que $\alpha, \alpha^2, \alpha^{2^2}, \ldots, \alpha^{2^r}$ son todas raíces del mismo polinomio, donde $\alpha^{2^r} = \alpha$ para algún $0 \leq r \leq t$. Supóngase ahora que $\gamma$ es un elemento primitivo de $F_{2^t}$ entonces $\alpha = \gamma^s$ para algún entero $s$ . Puesto que

$$\alpha^{2^r} = \alpha$$

$$\alpha^{2^r}\alpha^{-1} = 1$$

$$\gamma^{s2^r - s} = 1, \tag{2.3}$$

y además $\gamma$ es un elemento primitivo de orden $2^t - 1$ y por la ecuación (2,3) se cumple que $s2^r - s = k(2^t - 1)$, para algún entero $k$, así $s2^r \cong s\text{mód}(2^t - 1)$. Con base a las observaciones anteriores, se definen las clases ciclotómicas.

**Definición 32.** *La operación de múltiplicación por* 2*, divide a los enteros mód*$(2^t - 1)$ *en conjuntos disjuntos, llamados clases ciclotómicas mód*$(2^t - 1)$.

La clase ciclotómica que contiene a $s$ consiste de

$$C_s = \{s, 2s, 2^2 s, \ldots, 2^{r-1}s\}\text{mód}(2^t - 1),$$

donde $r$ es el entero positivo más pequeño tal que $s2^r \cong s\text{mód}(2^t - 1)$. Como la relación módulo es una relación de equivalencia entonces el conjunto formado por los $C_s$ es una partición del conjunto $\{0, 1, 2, \ldots, 2^t - 1\}$ de números enteros.

**Definición 33.** *Las n raíces de la unidad* $Z(C) = \{\alpha^i : i \in K\}$ *son llamadas ceros del código, donde* $K$ *es la unión de las clases ciclotómicas.*

El polinomio $z^n - 1$ se puede escribir

$$z^n - 1 = \prod_{i=0}^{n-1}(z - \alpha^i),$$

donde $\alpha$ denota una raíz n-ésima de la unidad.

**Definición 34.** *El polinimio minimal de* $\alpha^i \in F_{2^t}$ *es*

$$M^{(j)}(z) = \prod_{i \in C_s} (z - \alpha^i).$$

De las consideraciones anteriores se tiene que

$$z^n - 1 = \prod_s M^{(s)}(z).$$

**Ejemplo 6.**

Considérese $F_{2^3}$ un campo de extensión de $F_2$, construido a partir del polinomio irreducible $p(z) = 1 + z + z^3$. La tabla siguiente muestra los polinomio minimales sobre $F_2$ de cada elemento de $F_8$ y las clases ciclótómicas módulo siete.

| Raiz | Polinomio | clase |
|---|---|---|
| 0 | z | |
| 1 | $z + 1$ | $\{0\}$ |
| $\alpha, \alpha^2, \alpha^4$ | $z^3 + z + 1$ | $\{1, 2, 4\}$ |
| $\alpha^3, \alpha^5, \alpha^6$ | $z^3 + z^2 + 1$ | $\{3, 5, 6\}$ |

**Definición 35.** *Decimos que un código tiene* $r$ *raíces consecutivas si*

$$\left\{\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}\right\} \subseteq Z(C),$$

*para cierto natural* $b$*.*

**Teorema 26.** *Cota BCH*

Sea $C$ un código cíclico de longitud $n$ con polinomio generador $p(z)$ tal que para algunos enteros $b \geq 0$, $\delta \geq 1$

$$p(\alpha^b) = p(\alpha^{b+1}) = \cdots = p(\alpha^{b+\delta-2}) = 0.$$

Si el código tiene un arreglo de $\delta - 1$ potencias consecutivas de $\alpha$ como ceros, entonces la distancia mínima del código es al menos $\delta$.

**Demostración.**

Sea $c = (c_0, c_1, \ldots, c_{n-1})$ una palabra no nula de peso $w$ en $C$ y $c(z)$ su respectivo polinomio. Por hipótesis, el polinomio $p(z)$ tiene raíces que incluyen $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}$. Considérese la matriz

$$H = \begin{pmatrix} 1 & \alpha^b & \alpha^{2b} & \cdots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \cdots & \alpha^{(n-1)(b+1)} \\ & \vdots & & & \vdots \\ 1 & \alpha^{b+\delta-2} & \alpha^{2(b+\delta-2)} & \cdots & \alpha^{(n-1)(b+\delta-2)} \end{pmatrix}$$

la cual cumple que $Hc^T = \boldsymbol{0}$, para todo $c \in C$. En efecto, $c(z) = p(z)t(z)$ entonces

$$c(\alpha^b) = c(\alpha^{b+1}) = \cdots = c(\alpha^{b+\delta-2}) = 0.$$

Así $H$ es una matriz de paridad para el código $C$.

La idea de la prueba es mostrar que $w_c \geq \delta$. Supóngase que el vector $c$ tiene peso $w$, donde $w \leq \delta - 1$. Como existen $w$ componentes de $c$ diferentes de cero, considérese el conjunto $D = \{a_1, a_2, \ldots, a_w\}$, donde $c_i \neq 0$ si $i \in D$, luego la igualdad $Hc^T = \boldsymbol{0}$ permite dejar solamente en la matriz $H$ las columnas que corresponden a posiciones donde el vector $c$ tiene coordenadas no nulas, por tanto se puede escribir

$$\begin{pmatrix} \alpha^{a_1 b} & \alpha^{a_2 b} & \cdots & \alpha^{a_w b} \\ \alpha^{a_1(b+1)} & \alpha^{a_2(b+1)} & \cdots & \alpha^{a_w(b+1)} \\ \vdots & & & \vdots \\ \alpha^{a_1(b+w-1)} & \alpha^{a_2(b+w-1)} & \cdots & \alpha^{a_w(b+w-1)} \end{pmatrix} \begin{pmatrix} c_{a_1} \\ c_{a_2} \\ \vdots \\ c_{a_w} \end{pmatrix} = \boldsymbol{0}.$$

Puesto que el vector $c$ es no nulo, $H$ es una matriz singular y por tanto $detH = 0$. Pero $detH = \alpha^{(a_1+\cdots+a_w)b} detV$ donde $V$ es la matriz de Vandermonde

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{a_1} & \alpha^{a_2} & \cdots & \alpha^{a_w} \\ \vdots & & & \vdots \\ \alpha^{a_1(w-1)} & \alpha^{a_2(w-1)} & \cdots & \alpha^{a_w(w-1)} \end{pmatrix}.$$

Como todos los $\alpha^{a_j}$, $1 \leq j \leq w$ son distintos, el determinante de Vandermonde es distinto de cero, lo cual contradice el hecho de que el determinante de $H$ sea nulo. Por tanto el peso

$w$ de $c$ es mayor que $\delta - 1$, así para todo $c \in C$, $w(c) \geq \delta$, luego $w_c \geq \delta$. Por propiedades de los códigos lineales se tiene que $w_c = d_c$. Por tanto la distancia mínima del código es al menos $\delta$. $\nabla$

## 2.5. Códigos BCH

Hemos visto que los códigos cíclicos pueden ser descritos en términos de un polinomio generador $p(z)$. En esta sección se estudia otra forma, donde los códigos cíclicos son descritos por medio de las raíces de un polinomio generador, pertenecientes a algún campo de extensión. Si $\alpha \in F_{2^t}$ es raíz de $p(z)$, se cumple para todo $g(z)$ en el código, que $g(z) = p(z)r(z)$, luego $g(\alpha) = 0$, de igual forma si algún polinomio $g(z) \in R_n$, cuyas raíces incluyen los ceros de $p(z)$ tendrá a $p(z)$ como factor, luego $g(z)$ pertenece al código. Los códigos BCH son una familia de códigos diseñados para corregir múltiples errores y están definidos en términos de las raíces de sus polinomios generadores.

**Corolario 8.** *Un código cíclico de longitud $n$ con ceros $\alpha^b, \alpha^{b+r}, \alpha^{b+2r}, \ldots, \alpha^{b+(\delta-2)r}$ donde $r$ y $n$ son primos relativos tiene distancia mínima a lo menos $\delta$.*

**Demostración.**
Sea $\beta = \alpha^r$. Como $r$ y $n$ son primos relativos, $\beta$ es también una raíz n-ésima primitiva de la unidad, así $\alpha^b = \beta^t$ para algún $t$, además

$$\begin{array}{|rl|}\hline \beta^t = & \alpha^b \\ \beta^{t+1} = & \alpha^{b+r} \\ \beta^{t+2} = & \alpha^{b+2r} \\ \vdots & \\ \beta^{t+2} = & \alpha^{b+2r} \\ \hline \end{array}$$

y el código tiene los siguientes $\delta - 1$ ceros consecutivos $\beta^t, \beta^{t+1}, \beta^{t+2} \cdots, \beta^{t+\delta-2}$. Al reemplazar $\beta$ por $\alpha$ en la cota BCH se tiene el resultado del teorema. $\nabla$
Esta idea nos permite definir códigos cuya distancia mínima está acotada inferiormente.

**Definición 36.** *Un código de longitud n sobre $F_2$ es un código BCH con distancia asignada $\delta$, si para algún entero $b \geq 0$, su polinomio generador es*

$$p(z) = mcm[M^b(z), M^{b+1}(z), \ldots, M^{b+\delta-2}(z)],$$

*donde $M^b(z), M^{b+1}(z), \ldots, M^{b+\delta-2}(z)$ son los polinomios minimales de $\alpha^i$ respectivamente. $p(z)$ es el polinomio mónico de menor grado sobre $F_2$ que tiene como raíces $\alpha^i$ donde $b \leq i \leq b+\delta-2$.*

Además $c$ está en el código si y sólo si $c(\alpha^i) = 0$ para todo $b \leq i \leq b+\delta-2$. Así, el código tiene un arreglo de $\delta-1$ potencias consecutivas de $\alpha$ como raíces. Por el corolario anterior, la distancia mínima de un código BCH es mayor o igual que la distancia asignada $\delta$.

**La matriz de chequeo de paridad de un código BCH tiene la forma**

$$H = \begin{pmatrix} 1 & \alpha^b & \alpha^{2b} & \cdots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \cdots & \alpha^{(n-1)(b+1)} \\ & \vdots & & & \vdots \\ 1 & \alpha^{b+\delta-2} & \alpha^{2(b+\delta-2)} & \cdots & \alpha^{(n-1)(b+\delta-2)} \end{pmatrix}$$

**Comentarios**

- Si $b = 1$ el código BCH es llamado *narrow sense*.
- Si $n = 2^m - 1$ el código es llamado primitivo, donde $\alpha$ es un elemento primitivo del campo $F_{2^m}$.
- Si algún $\alpha^i$ es un cero en el código entonces también lo son todas los $\alpha^l$, para $l$ en la clase ciclotómica $C_i$.

## 2.6. Aplicaciones de los códigos lineales

Los datos digitales mueven actualmente la sociedad. La enorme cantidad de información que se incorpora cada día a la Internet, las tarjetas de crédito, los informes y registros bancarios, el campo de las comunicaciones telefónicas, la administración fiscal, los datos generados en el mundo laboral, las imágenes, la música y muchos más. En todos estos campos se genera gran cantidad de información, la cual es necesario enviar de un sitio a otro en forma segura. El intercambio de información ha dejado de ser algo simplemente implícito en los sistemas actuales y pasó a reclamar un papel central y protagonista, en los planes y estrategias informáticas en el ámbito doméstico y empresarial.
En los últimos años, la codificación ha adquirido un alto grado de importancia en la vida cotidiana y científica debido a las múltiples posibilidades de mejorar la calidad y seguridad en las comunicaciones, por esta razón, los códigos correctores de errores se aplican a una gran variedad de sistemas de información, entre ellos se destacan

- Sistemas de comunicación: vía satélite, microondas (radio y televisión), sistemas de teletexto, videoconferencia, internet.
- Sistemas informáticos: circuitos lógicos, memorias de semiconductores, discos flexibles, discos duros, memorias flash, discos de lectura óptica (CD-ROM,CD-RW).
- Sistemas de audio y video: sonido digital (CD), video digital (DVD).

El código de Hamming es utilizado para identificar y corregir un bit erróneo en una palabra codificada, actualmente son utilizados en el proceso de lectura y escritura de una memoria RAM.
Los códigos cíclicos se aplican en telefonía móvil, telefonía para comunicación de voz IP, en especial han sido utilizados por la NASA para transmisión de información en las misiones espaciales, como fueron la sonda Galileo a Júpiter en 1989, la sonda Magallanes a Venus ese mismo año o la sonda Ulises al Sol en 1990).
Los códigos BCH son utilizados actualmente para codificar y grabar los datos en discos compactos y mejorar la confiabilidad de estos.

Además en los nuevos campos de investigación encontramos que las ideas de Shannon se aplican en medicina, en teorías evolutivas, biología, etc. Por ejemplo, se ha planteado la posibilidad de medir qué tanta información puede absorber, procesar y emitir un cerebro humano. Experiencias llevadas a cabo demuestran sorpresivamente que el tiempo de respuesta a cierto tipo de estímulo es una función logarítmica de la cantidad de fuentes de estímulos utilizadas. Otras investigaciones utilizan estas herramientas para medir la capacidad de un nervio, o para medir la capacidad de almacenamiento de información de una molécula de ADN. Algunos estudios sugieren la posibibilidad de que estas moléculas puedan utilizar protocolos correctores de errores. Se ha propuesto que el paso de los años produce en el ADN cierto tipo de averías que afectan su capacidad de transmitir la información vital, es decir, con el envejecimiento se incrementa el ruido en nuestros organismos.

## 2.7. Comentarios

1. Este trabajo se realizó con base a un campo $F_2$, pero la teoría expuesta se puede extender a campos $F_q$, de esta forma se definen códigos lineales con condiciones para establecer sus parámetros. Tal es el caso de los códigos Reed-Solomon, una subfamilia importante de los BCH, con la propiedad especial que su longitud es $n = q - 1$. Construir códigos Reed-Solomon con longitud $n = 1$ no es de gran utilidad, dado que no son eficientes, por lo cual, estos códigos necesitan ser especificados sobre un campo con $q > 2$. Esta es la razón por la cual no son analizados en este texto.

2. Los algoritmos basados en codificaciones cíclicas, brindan posibilidades a los investigadores de encontrar mejores códigos para sus aplicaciones, gracias a su estructura matemática la cual permite manipular polinomios de grados muy grandes, permitiendo aumentar la longitud del código, logrando así mayor capacidad de codificar una gran cantidad de información, de igual forma aumenta el número de raíces consecutivas del código y por tanto la distancia mínima, se tiene entonces una mayor capacidad de corregir errores en forma óptima.

3. Se utilizan otras herramientas de la matemática para definir códigos, tales como la teoría de grafos usada para la representación de los códigos convolucionales, también se han construido códigos utilizando geometrías algebraicas, estos están definidos sobre curvas en un espacio. En proyectos de investigaciones recientes se construyen nuevos códigos haciendo uso de conjuntos $B_h$.

# Bibliografía


[1] A. Montenegro. *Códigos de detección de error*, Universidad del Cauca, 2002.

[2] D.R. Shier, K.T. *Wallenius Applied mathematical modeling a multidisciplinary approach*, Chapman Hall–CRC, 1999.

[3] E.F. Assmus Jr, J.D Key. *Designs and their codes*, Cambridge University Press, 1992.

[4] F.J. Macwillians and N.J.A. Slone *The theory of Error-Correcting Codes*, North-Holland Matematical library, 2006.

[5] J.H. Van Lint. *Introduction to Coding Theory*. Second Edition, Springer–Verlag, 1991.

[6] J.I. Hall *Notes on coding theory* , Michigan State University, 2003

[7] J.L. Tábara. *Códigos Correctores de error*, 2002.

[8] L. Couch. *Sistemas de comunicación digital y analógica*, Prentice Hall 1999.

[9] R.L. Harald. *Niederreiter Finite Fields*, Cambridge University Press, 1997.

[10] V. Pless. *Introduction to the Theory of Error–Correcting Codes*. Second Edition, WILEY, 1989.

[11] W. Cary Huffman and Vera pless *Fundamentals of Error-Correcting Codes*, Cambridge University Press, 2003.