CÓDIGOS LDPC: UN ACERCAMIENTO

TESIS QUE PRESENTA

LAURA ALEJANDRA GÓMEZ TEXCO

PARA OBTENER EL TÍTULO DE
LICENCIADA EN MATEMÁTICAS APLICADAS

ASESOR DE TESIS: DR. CARLOS ALBERTO LÓPEZ ANDRADE

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA ⟨BUAP⟩
Facultad de Ciencias Físico Matemáticas ⟨FCFM⟩
http://www.fcfm.buap.mx/
Noviembre 2017

WEBSITE:
http://www.fcfm.buap.mx/docencia/tesis/licenciatura-ma

E-MAIL:
gt.ale.laus@gmail.com

DEDICATORIA

Tú eres el milagro más grande del mundo.

– Og Mandino –

Mi trabajo de tesis va dedicado a aquella chica que decidió estudiar la Licenciatura en Matemáticas Aplicadas, pues sin esa decisión yo no hubiera escrito estas líneas. A esa chica que ha ido (y seguirá) conociendo poco a poco lo que son las matemáticas. A esa chica perseverante, a la que nunca he visto darse por vencida y mucho menos que algo o alguien la derrote. Para mí, ella es un ser espiritual y humano completo, que nunca se pierde de algún detalle de mi vida, ni de mis alegrías, ni mis tristezas, y mucho menos de esos instantes caóticos en donde debe tratar que lo complejo de la vida se vuelva simple. Dedicada a ese ser que amo y amaré esta vida y las próximas. Dedicada a mí.

Laura Alejandra G. T.

—Es necesario que digas lo que piensas.
—Eso es precisamente lo que estoy haciendo –respondió Alicia–, o al menos pienso lo que digo, que es lo mismo, ¿no?
—De ningún modo –dijo el Sombrerero–...
—Y también podrías decir –dijo ahora la Liebre de Marzo–: "me gusta lo que tengo", como si fuera lo mismo que "tengo lo que me gusta".

– Alicia en el país de las maravillas - Lewis Carroll –

AGRADECIMIENTOS

Uno puede devolver un préstamo de oro, pero está en
deuda de por vida con aquellos que son amables.

– Proverbio –

Es muy grato para mi reconocer y dar mérito de este logro a seis personitas muy fundamentales en mi vida, pues sin ellos yo no hubiera podido llegar hasta este punto. Ellos me han llenado el alma de los más preciados sentimientos y valores; y siempre me han brindado el mayor apoyo sin esperar algo a cambio más que mi bienestar y felicidad. Quiero agradecerles muy en especial e infinitamente con todo mi ser, mi cariño, respeto y amor, pues este también es un logro de ustedes, a mi mamá Micaela Texco Velázquez (una mujer increíblemente maravillosa), mi papá Manuel Gómez Juárez (un hombre honesto y trabajador), mis extraordinarios hermanos, Luis Gerardo (Ñero), Miguel (Miguelas), Manuel David (Davichón), y mi hermosa hermanita Jhoanita (mi inspiración), gracias por compartir su vida conmigo, creer y confiar en mí, son parte de lo más valioso que la vida me ha dado, los amo.

Agradezco a todos mis profesores y profesoras por compartirme sus conocimientos, su sabiduría, y trato, pues todo ello ha contribuido para mi formación profesional y humana. En particular, quiero agradecer a mis sinodales por tomarse el tiempo para leer mi tesis y contribuir a mejorarla. Y en especial, quiero agradecer a mi asesor de tesis, el Dr. Carlos Alberto López Andrade, ya que la realización de este proyecto fue un trabajo conjuntamente de ambos, y puedo decir que a parte de ser un gran matemático, como persona es un ser excepcional.

Agradezco a todas mis amigas y amigos, porque con ustedes me divertí (literal), socialicé y he vivido grandes aventuras, y sobre todo he conocido diversos puntos de vista y personalidades.

Y antes de concluir, también quiero agradecer a un ser muy valioso y raro en mi vida. De quien he y con quien he aprendido muchas cosas, que me llena la mente de curiosidades, y el alma de felicidad, de dicha y amor. Pues me ayuda a vencer mis miedos, fortalecer mis virtudes, y a continuar en eso que me apasiona, (sé que cuando lo leas sabrás que eres tú).

Finalmente quiero agradecer al universo, a la vida, y a Dios, por esta increíble oportunidad de vivir un día más; y agradecer a todo ser humano que se ha cruzado en mi camino, porque creo que todos han contribuido a crear quien soy.

No me queda más que decir... ¡Gracias!

Laura Alejandra G. T.

INTRODUCCIÓN

Lo último que uno sabe, es por dónde empezar.

– Blaise Pascal –

¿Para qué estudiar a los códigos LDPC? Para saber que LDPC quiere decir Chequeo de Paridad de Baja Densidad (Low Density Parity Check por sus siglas en inglés). Que estos códigos fueron inventados por Gallager olvidados un cierto tiempo y redescubriertos por Mackay y Neal, para saber que los códigos LDPC son una clase de códigos de bloque lineal, que su característica esencial es tener una matriz de chequeo de paridad de baja densidad, que tienen una representación gráfica bipartita conocida como grafo de Tanner, y que su decodificación es mediante algoritmos iterativos de paso de mensaje; y además, que se acercan bastante bien al límite de capacidad establecido en el teorema de Shannon. En sí, para saber que los códigos LDPC tienen una estructura matemáticamente interesante. Ésta podría ser una razón, pero la realidad es que los códigos LDPC son estudiados para dar solución a un problema: al problema central de la teoría de la comunicación, cuyo objetivo es lograr construir un sistema de codificación y de decodificación que haga posible la comunicación confiable, eficaz y segura; y además, que sea aceptable para las aplicaciones prácticas, es decir, requerimos buenos códigos detectores-correctores de errores capaces de ser implementados para recuperar la palabra-código original produciendo así, una buena transmisión de la información. Entonces, el interés por los códigos LDPC se debe a que son códigos que pueden recuperar la información original ante la presencia de grandes cantidades de ruido y tienen baja complejidad para su implementación en una gran variedad de medios de comunicación. Los códigos LDPC no son los únicos que intentan dar solución al problema de la comunicación pues existen otros códigos.

En este proyecto de tesis, presentamos un acercamiento al estudio de la teoría fundamental de los códigos LDPC, aunque el objetivo principal es realizar un análisis detallado de tres teoremas que establecen una cota para la distancia mínima del código. Para llevar a cabo esto, debemos conocer un poco sobre la teoría de códigos, teoría de grafos (o gráficas), teoría de la información, y el álgebra lineal, esto se revisará y estudiará en los capítulos 2 y 3; en el capítulo 4, nos centraremos en los códigos LDPC; y para contextualizar a los códigos LDPC, en el capítulo 1, hablaremos un poco sobre el problema de la teoría de la comunicación. El trabajo de la tesis está dividido en cuatro capítulos, los cuales describimos a continuación.

En el capítulo uno contextualizamos a los códigos. Abordamos la problemática de la comunicación, dando un panorama acerca de cómo poder dar solución a ese problema, a partir del uso de códigos como el código de repetición $\mathcal{R}_3$ y el $[7,4]$-código de Hamming. Para ambos códigos analizamos su codificación y decodificación, y en el caso del $[7,4]$-código de Hamming vemos que estos procesos puede analizarse mediante una representación gráfica, para finalmente dar un primer enfoque de que hay grafos asociados a códigos y viceversa.

En el capítulo dos definimos los objetos a usar. Revisamos los conceptos básicos de un código, la estructura de los códigos lineales, el síndrome de decodificación, lo que es un código detector-corrector de errores, así como un teorema importante que nos permite saber cuantos errores es capaz de corregir un código conociendo su distancia mínima. Después, enunciamos el Teorema de codificación de canal con ruido de Shannon para darnos una idea de lo que establece. También, revisamos conceptos básicos de grafos, la matriz de adyacencia e incidencia, así como el concepto de grafo bipartito; además, estudiamos el concepto de expansión de un grafo bipartito. Y finalmente estudiamos la relación entre los códigos lineales y los grafos bipartitos.

En el capítulo tres revisamos los resultados a ocupar. Analizamos algunos conceptos y resultados de álgebra lineal, como valores propios, vectores propios, diagonalización y ortogonalización. También, estudiamos el Teorema de descomposición espectral. Y finalmente el espectro de un grafo no dirigido.

Por último, en el capítulo cuatro estudiamos a los códigos LDPC. El objetivo es estudiar el concepto de un código LDCP regular e irregular, luego, caracterizamos la irregularidad de un código LDPC a partir del cálculo de las distribuciones de grado. Y además, realizamos un análisis detallado sobre las demostraciones de tres teoremas que establecen una cota para la distancia mínima del código. Pero antes de eso, damos una introducción sobre el panorama general de lo que abarca el estudio, la implementación y las aplicaciones de los códigos LDPC.

Esperamos que esta tesis "Códigos LDPC: un acercamiento" sea un buen referente para aquellos lectores que estén interesados en introducirse al tema, pues este trabajo está basado en diversos escritos sobre códigos LDPC.

ÍNDICE GENERAL

1 PRELIMINARES

Tres cosas deberían ser consideradas:
los problemas, los teoremas, y las aplicaciones.

– Gottfried Wilhelm Leibniz –

En este capítulo hablaremos sobre la problemática de la comunicación digital. El proceso de comunicación puede describirse de la siguiente manera: se tiene un emisor el cual desea transmitir, almacenar o enviar algún tipo de información; esta información es transmitida a través de un medio o canal con ruido, esto provocará errores en la información produciéndose así una información modificada, de tal forma la información que recibe el receptor será una información diferente de la original. En sí, esto no es deseable para ningún sistema que realice este proceso de comunicación, por lo que poco a poco fueron surgiendo diferentes maneras de solucionarlo, y una de las cuales fue la implementación de los códigos detectores-correctores de errores.

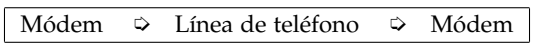
El texto base revisado para el desarrollo de este capítulo fue [Mac03].

1.1 EL PROBLEMA DE LA COMUNICACIÓN

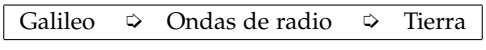
Hoy en día, las redes de comunicación comparten el mismo principio fundamental de operación, ya sea que se trate de paquetes de datos a través de la Internet, o señales en una red telefónica, o vía satélite, y también en el almacenamiento digital de datos, la información se transmite de la misma manera, a través de un canal de comunicación ruidoso. El ruido provoca errores que impiden obtener una buena transmisión de la información, produciéndose pérdida de información, falta de seguridad y fiabilidad en la red o sistema. De esta manera, una de las interrogantes más esenciales de esta situación es: ¿cómo podemos lograr una comunicación perfecta sobre algo imperfecto, es decir, sobre un canal de comunicación con ruido? Durante el desarrollo de esta sección obtendremos una idea de cómo es posible lograr una comunicación perfecta a través de canales de comunicación imperfectos.

Algunos ejemplos de canales de comunicación ruidosos:

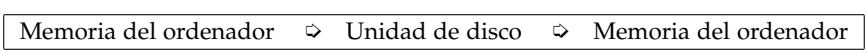
Ejemplo 1.1 Línea telefónica, sobre los cuales dos módems comunican información digital. Los problemas que enfrentan es que: o bien sufre de conversaciones cruzadas, o ruido en la señal de transmisión, o distorsión de la línea.



Ejemplo 1.2 Enlace de comunicación de radio, como el enlace que existe desde la nave espacial moviéndose en la órbita de Júpiter, Galileo, a la Tierra. Esta transmisión a parte de ser una señal muy débil, el canal por el que viaja recibe radiación terrícola y cósmica.



Ejemplo 1.3 Unidad de disco, escribir un dígito binario en el material magnético y fallar al momento de leer la memoria del dígito almacenado.



El último ejemplo muestra que la comunicación no tiene que involucrar información que esté yendo de un lugar a otro.

En todos estos casos, si transferimos datos, por ejemplo, una cadena de bits, sobre un canal, hay cierta probabilidad de que el mensaje recibido no sea idéntico al mensaje transmitido. Lo ideal sería tener un canal de comunicación con probabilidad de error muy cercana a cero. La Figura 1 describe la problemática del proceso de comunicación de la forma más simple.

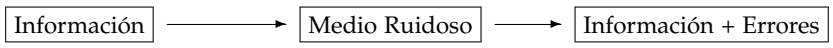


Figura 1: Proceso de comunicación, esquema sin código, en la que se almacena o transmite información la cual atraviesa por un medio ruidoso que modificará la información.

Pasemos a una estructura un poco más rigurosa, usando terminología matemática. Consideremos una unidad de disco ruidoso que transmite cada bit correctamente con probabilidad $1-f$ (probabilidad de éxito) e incorrectamente con probabilidad f (probabilidad de fracaso). Este modelo de canal de comunicación es conocido como *canal simétrico binario (CSB)*. Definamos lo siguiente: x mensaje transmitido, y mensaje recibido, f probabilidad de que el bit sea intercambiado.

$$P(y_1 = 0|x_1 = 0) = 1-f, \qquad P(y_2 = 1|x_1 = 0) = f;$$
$$P(y_1 = 0|x_2 = 1) = f, \qquad P(y_2 = 1|x_2 = 1) = 1-f;$$
$$\text{y} \quad \sum_{j=1}^{2} P(y_j|x_i) = 1 \quad \text{para cada } i = 1, 2.$$

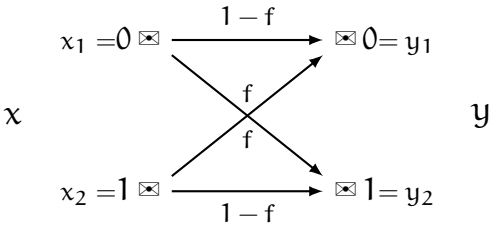


Figura 2: Canal simétrico binario (CSB).

La Figura 2 representa el modelo del CSB. Describiremos su función de la siguiente manera: cada *mensaje transmitido* x y *mensaje recibido* y está constituido de una secuencia de bits, en este caso de los dígitos binarios 0 y 1, supóngase que el primer bit transmitido es 0, el cual al pasar por el canal tiene probabilidad de éxito $1-f$, es decir, la probabilidad $1-f$ es la capacidad que tiene este bit 0 de ser enviado correctamente, así el bit recibido sería nuevamente 0, pero no se debe descartar la otra opción, donde dado que el primer bit transmitido es 0 y que el bit recibido sea 1, diferente del bit original, debido a la probabilidad f de que el bit sea intercambiado. Supóngase ahora que el segundo bit transmitido es 1, éste de igual forma cuenta con dos opciones, la probabilidad $1-f$ que indica la capacidad de enviar el bit sin errores y éste sea 1, el mismo que el original, o la probabilidad f que causaría que el bit recibido sea 0 distinto del original. Debido a que tanto el bit 0 y 1 tienen dos opciones con las mismas probabilidades de éxito y fracaso, ésta es la razón de que se cumple con la simetría del modelo.

Lo descrito en el párrafo anterior, podemos resumirlo en la *matriz del canal* M_C, la cual está definida como la matriz cuyas componentes son las probabilidades de transición en el canal, en este caso para el canal simétrico binario la matriz del canal es:

$$M_C = \begin{matrix} \\ 0 \\ 1 \end{matrix} \begin{matrix} \begin{matrix} 0 & \quad 1 \end{matrix} \\ \begin{bmatrix} P(y_1|x_1) & P(y_2|x_1) \\ P(y_1|x_2) & P(y_2|x_2) \end{bmatrix} \end{matrix} = \begin{bmatrix} 1-f & f \\ f & 1-f \end{bmatrix} \qquad \textit{Matriz del canal simétrico binario.}$$

Para responder a la pregunta planteada en el primer párrafo tenemos dos tipos de soluciones: la solución física y la solución del sistema, las cuales platicaremos a continuación.

1.1.1 La solución física

En el apartado de la solución física tenemos las soluciones que en realidad sólo tratan de mejorar las características físicas del canal de comunicación para reducir la probabilidad de error. Para lo cual debemos:

1) Usar más componentes confiables en el sistema de circuitos.

2) Evacuar el aire desde la cerradura del disco así como eliminar la turbulencia que perturba al leer la cabecera desde la pista.

3) Usar un parche magnético más grande para representar cada bit.

4) Utilizar señales muy poderosas o enfriar el sistema de circuitos en orden para reducir el ruido termal.

Sin embargo, estas modificaciones físicas incrementan los costos del canal de comunicación.

1.1.2 La solución de sistema

La teoría de la información y la teoría de la codificación ofrecen una alternativa distinta, ya que con la práctica de estas teorías, es posible detectar y corregir los errores introducidos por el canal al mensaje original, obteniendo así el mensaje de salida con la menor cantidad de errores.

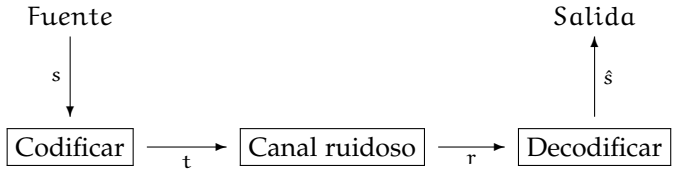


Figura 3: Proceso de comunicación, esquema con código.

La Figura 3 es un esquema sobre la solución de sistema añadida al proceso de comunicación, logrando así una comunicación confiable sobre un canal ruidoso. Observemos que añadimos una codificación antes del canal y una decodificación después. El codificador codifica el *mensaje fuente* s obteniendo un *mensaje transmitido* t, al realizar la codificación lo que realmente ocurre es que se agrega redundancia al mensaje original de algún modo para reducir la probabilidad de fracaso y mejorar la probabilidad de éxito. El canal añade ruido al mensaje transmitido, produciendo el *mensaje recibido* r, donde el mensaje recibido r es el mensaje transmitido t con posibles errores. Posteriormente, el decodificador decodifica el mensaje recibido r para obtener un *mensaje de salida* $\hat{s}$, esta decodificación se realiza para tratar de detectar y corregir los errores, y así obtener un mensaje de salida $\hat{s}$ el cual deseamos que sea lo suficientemente idéntico al mensaje original s.

Esta solución de sistema hace más confiable la comunicación, pero aún se tienen limitantes, tanto los desarrollos tecnológicos como los avances teóricos.

1.2 CÓDIGOS DETECTORES-CORRECTORES DE ERRORES PARA EL CANAL SIMÉTRICO BINARIO

¿Cuál es el camino más simple para añadir redundancia útil en una transmisión?
Reglas del juego:

- Ser capaces de detectar y corregir errores.
- La retransmisión no es una opción.
- Tener una sola oportunidad para codificar, transmitir y decodificar.

1.2.1 Códigos de repetición

En un *código de repetición* la idea es repetir cada símbolo de información un número predeterminado de veces.

Definición 1.1. $\mathcal{R}_m \subseteq \mathbb{F}_2^m$ *es un código de repetición de longitud m, si dado* $c = (c_0, c_1, \ldots, c_{m-1}) \in \mathcal{R}_m$ *se cumple que* $c_0 = c_1 = \cdots = c_{m-1}$.

Ejemplo 1.4 El código de repetición de longitud 3, $\mathcal{R}_3$, se muestra en la Tabla 1.

Secuencia fuente	Secuencia transmitida
s	t
0	000
1	111

Tabla 1: Código de repetición $\mathcal{R}_3$.

1.2.2 Codificación: Uso del código de repetición $\mathcal{R}_3$

El ejemplo siguiente nos permitirá observar el proceso de comunicación implementando códigos detectores-correctores de errores, que se analizó en la Figura 3.

Ejemplo 1.5 Sea s el mensaje o vector fuente, t el mensaje o vector transmitido, e el vector ruido, r el mensaje o vector recibido, $\hat{s}$ el mensaje o vector de salida, y f el nivel de ruido en el canal (CSB) que es la probabilidad de que el bit sea intercambiado. Imaginemos que transmitimos el mensaje fuente

$$s = 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0$$

sobre un canal simétrico binario con nivel de ruido $f = 0.1$ usando el código de repetición $\mathcal{R}_3$.

s	0	0	1	0	1	1	0
t	000	000	111	000	111	111	000
e	000	001	000	000	101	000	000
r	000	001	111	000	010	111	000

Tabla 2: Un ejemplo de transmisión usando $\mathcal{R}_3$.

En la Tabla 2, observamos que el primer renglón es el mensaje fuente $s = 0010110$, el segundo renglón es el mensaje transmitido $t = 000000111000111111000$ en donde cada bit de s fue codificado de acuerdo al código de repetición $\mathcal{R}_3$, luego el tercer renglón es el vector ruido $e = 000001000000101000000$ que se añade al canal, y finalmente obtenemos el mensaje recibido $r = 000001111000010111000$ que es el resultado de la suma de t y e módulo 2, es decir, el mensaje recibido es $r = t + e$ (mód 2).

1.2.3 Decodificación: ¿Cómo decodificamos el vector recibido r?

La idea central del algoritmo óptimo de decodificación es, observar a los tres bits recibidos al mismo tiempo y elegir uno por mayoría de votos. La decisión óptima de decodificación (óptima en el sentido de tener la menor probabilidad de estar equivocado) es encontrar cuál valor de s es más probable, dado r. Consideremos la decodificación de un solo bit s (aquí s denota un solo bit del vector fuente), el cual fue codificado como $t(s) = t_1(s)t_2(s)t_3(s)$ por el código de repetición $\mathcal{R}_3$ y dados los tres bit recibidos $r = r_1r_2r_3$, queremos hallar el valor de s más probable dado r, es decir:

$$P(s|r).$$

Ahora bien, $P(s|r)$ es la probabilidad condicional definida como:

$$P(s|r) = \frac{P(r \cap s)}{P(r)} \qquad \text{Probabilidad Condicional.}$$

Al multiplicar por la unidad tenemos:

$$\begin{aligned} P(s|r) &= \frac{P(r \cap s)}{P(r)} \cdot \frac{P(s)}{P(s)} \\ &= \frac{P(r \cap s)}{P(s)} \cdot \frac{P(s)}{P(r)} \\ &= P(r|s) \cdot \frac{P(s)}{P(r)} \\ &= \frac{P(r|s)P(s)}{P(r)}. \end{aligned}$$

Obteniendo así, el resultado conocido como Teorema de Bayes:

$$P(s|r) = \frac{P(r|s)P(s)}{P(r)}. \tag{1.1}$$

No perdamos de vista que nuestro interés es calcular $P(s|r) = P(s|r_1r_2r_3)$. Para los valores de s solo hay dos alternativas, pues $s \in \{0, 1\}$, y usando el resultado (1.1) tenemos que:

$$P(s = 0|r_1r_2r_3) = \frac{P(r_1r_2r_3|s = 0)P(s = 0)}{P(r_1r_2r_3)}, \tag{1.2}$$

$$P(s = 1|r_1r_2r_3) = \frac{P(r_1r_2r_3|s = 1)P(s = 1)}{P(r_1r_2r_3)}. \tag{1.3}$$

Dichas probabilidades (1.2) y (1.3), están determinadas por los valores de $P(s)$, $P(r_1r_2r_3|s)$ y $P(r_1r_2r_3)$. Pero solo nos enfocaremos en calcular los valores de $P(s)$ y $P(r_1r_2r_3|s)$, y no necesitaremos calcular $P(r_1r_2r_3)$, pues al encontrar la decisión óptima de decodificación, supondremos que:

$$\hat{s} = \begin{cases} 0 & \text{si} \quad P(s = 0|r) > P(s = 1|r), \\ 1 & \text{si} \quad P(s = 1|r) > P(s = 0|r). \end{cases}$$

Notemos que $P(s)$ es un valor conocido, ya que como mencionamos anteriormente s solo toma dos valores y estos son equiprobables, así, $P(s) = \frac{1}{2}$ con $s \in \{0, 1\}$. Considerando esta observación, y tomando en cuenta las igualdades (1.2) y (1.3), las condiciones de la decisión óptima de decodificación se reducen a:

$$\hat{s} = \begin{cases} 0 & \text{si} \quad P(r|s = 0) > P(r|s = 1), \\ 1 & \text{si} \quad P(r|s = 1) > P(r|s = 0). \end{cases} \tag{1.4}$$

Asumiendo que estamos sobre un canal simétrico binario con nivel de ruido $f < 0.5$, calculemos $P(r|s)$:

$$\begin{aligned} P(r|s) &= P(r_1r_2r_3|s) \\ &= P(r_1r_2r_3|t(s)) \qquad (t(s) \text{ es la codificación de } s) \\ &= P(r_1r_2r_3|t_1(s)t_2(s)t_3(s)) \\ &= P(r_1|t_1(s)t_2(s)t_3(s)) \cdot P(r_2|t_1(s)t_2(s)t_3(s)) \cdot P(r_3|t_1(s)t_2(s)t_3(s)) \\ &\quad \text{(dado que } r_1, r_2, r_3 \text{ son independientes dos a dos)} \\ &= P(r_1|t_1(s)) \cdot P(r_2|t_2(s)) \cdot P(r_3|t_3(s)) \\ &\quad \text{(debido a que } r_i \text{ solo depende de } t_i(s),\ i = 1, 2, 3) \\ &= \prod_{i=1}^{N=3} P(r_i|t_i(s)), \end{aligned}$$

donde N es el número de bits transmitidos en el bloque, y:

$$P(r_i|t_i) = \begin{cases} 1 - f, & \text{si} \quad r_i = t_i, \\ f, & \text{si} \quad r_i \neq t_i. \end{cases}$$

Luego, considerando el cociente de probabilidades para las dos hipótesis de (1.4), tenemos:

$$\frac{P(r|s=1)}{P(r|s=0)} = \frac{\prod_{i=1}^{N=3} P(r_i|t_i(1))}{\prod_{i=1}^{N=3} P(r_i|t_i(0))} = \prod_{i=1}^{N=3} \frac{P(r_i|t_i(1))}{P(r_i|t_i(0))},$$

cada factor $\frac{P(r_i|t_i(1))}{P(r_i|t_i(0))}$ es igual a $\frac{1-f}{f}$ si $r_i = 1$ y $\frac{f}{1-f}$ si $r_i = 0$. Así, definimos el cociente:

$$\gamma = \frac{1-f}{f}.$$

Con lo que, dado $f < 0.5$ podemos decidir que valor otorgarle a $\hat{s}$ el bit de salida, tomando en cuenta que:

- Si $\frac{P(r|s=1)}{P(r|s=0)} > 1$ eso implica que $P(r|s=1) > P(r|s=0)$ entonces $P(s=1|r) > P(s=0|r)$ y así $\hat{s} = 1$.
- Si $\frac{P(r|s=1)}{P(r|s=0)} < 1$ eso implica que $P(r|s=0) > P(r|s=1)$ entonces $P(s=0|r) > P(s=1|r)$ y así $\hat{s} = 0$.

Por lo tanto, la decodificación por mayoría de votos es la decodificación óptima si asumimos que el canal es un canal simétrico binario y que los dos posibles mensajes fuentes 0 y 1 son equiprobables.

Secuencia recibida r	Cociente de probabilidades $\frac{P(r\|s=1)}{P(r\|s=0)}$	Secuencia decodificada $\hat{s}$
000	$\gamma^{-3} < 1$	0
001	$\gamma^{-1} < 1$	0
010	$\gamma^{-1} < 1$	0
100	$\gamma^{-1} < 1$	0
101	$\gamma^{1} > 1$	1
110	$\gamma^{1} > 1$	1
011	$\gamma^{1} > 1$	1
111	$\gamma^{3} > 1$	1

Tabla 3: Resultados del algoritmo de decodificación por mayoría de votos para $\mathcal{R}_3$.

La Tabla 3 muestra los resultados de la decodificación aplicada a las posibles secuencias recibidas de $\mathcal{R}_3$.

s	0	0	1	0	1	1	0
t	000	000	111	000	111	111	000
e	000	001	000	000	101	000	000
r	000	001	111	000	010	111	000
$\hat{s}$	0	0	1	0	0	1	0
Errores corregidos		✓					
Errores no corregidos					✗		

Tabla 4: Resultado de la decodificación al vector recibido r.

Apliquemos el algoritmo de decodificación al vector recibido r de la Tabla 2. Para el primer bloque 000, los tres primeros bits recibidos son todos 0, así al decodificar este bloque obtenemos 0, aquí es notable que no hay error que corregir. En el segundo bloque 001, hay dos 0's y un 1, así al decodificar obtenemos 0, en este caso corregimos el error. Para el tercer bloque 111, todos los bits son 1, decodificando obtenemos 1, y nuevamente no había error. Es observable que para los bloques de bits cuarto, sexto y séptimo no hubo problema ni error de correción. Sin embargo, para el quinto bloque 010, había dos bits 0 y un bit 1, al decodificar obtenemos 0, pero es claro que no corregimos el error, pues el bit fuente en realidad fue 1. Veamos como resultó nuestra decodificación, Tabla 4.

Observemos, que dicho algoritmo es capaz de corregir un solo error, y en donde haya dos errores o más esta decodificación no tiene la capacidad de corregirlos, es decir, el error no será corregido y de esta forma el algoritmo fallará.

La probabilidad de error es determinada por las probabilidades de que dos o más bits en un bloque de tres sean intercambiados, la probabilidad de intercambio de dos bits de tres es f^2 y la probabilidad de intercambio de los tres bits es f^3. Al considerar el caso para el código de repetición $\mathcal{R}_3$, tenemos 3 combinaciones de dos de tres bits que pudieran ser intercambiados para los que siempre sucede que un bit sigue siendo el mismo con probabilidad $(1-f)$, por lo que la probabilidad de error está dada por $3f^2(1-f)+f^3$. En nuestro ejemplo, tenemos el canal simétrico binario con un nivel de ruido de $f = 0.1$, así al implementar el código de repetición $\mathcal{R}_3$ se tiene una probabilidad de error después de la decodificación, $P_b \simeq 0.03$ por bit. Generalizando, la probabilidad de error para el código $\mathcal{R}_m$, el código de m repeticiones, donde m es impar, está dado por:

$$P_b = \sum_{i=\frac{m+1}{2}}^{m} \binom{m}{i} f^i (1-f)^{m-i}$$

Por lo tanto, la probabilidad de error es reducida al usar algún código de repetición $\mathcal{R}_m$ en un canal simétrico binario con nivel de ruido f, aunque no se asegura por completo que el mensaje de salida $\hat{s}$ sea identico al mensaje fuente s.

1.2.4 Un ejemplo implementando el código de repetición $\mathcal{R}_3$ y el algoritmo de decodificación por mayoría de votos

Considérese una imagen binarizada como la Figura 4, ésta será codificada con el código de repetición $\mathcal{R}_3$, para luego añadirle ruido al canal, finalmente la decodificamos usando el algoritmo por mayoría de votos. Al final del programa realizado en el entorno de *Processing 2* [FR01], Código 1, mostramos los resultados de la ejecución, las imágenes de salida, Tabla 5.

Figura 4: Imagen fuente.

Código 1: Programa que muestra el resultado de una codificación y decodificación.

```
PImage img;
int w, h;
void setup()
{
  img=loadImage("img_fuente.png");
  img.loadPixels();
  w=img.width;
  h=img.height;
  size(w,h);
```

```
  code();
}

int i=0, j=0;
float aux, P, p1=1.0, p2=1.0, p3=1.0;
float s; // bit fuente
float t1, t2, t3; // bloque de bits transmitidos, t = t1t2t3
float r1, r2, r3; // bloque de bits recibidos, r = r1r2r3
float b; // bit salida
float f = 0.125; // nivel de ruido en el canal
float R; // cociente de probabilidades
void code()
{
  println(w*h);
  while(i<(w*h))
  {
    s = red(img.pixels[i]); // bit fuente

    if(s==255) // codificacion de bit fuente s (codigo de repeticion R3) *
    {
      t1 = 1;
      t2 = 1;
      t3 = 1;
    }//blanco
    else
    {
      t1 = 0;
      t2 = 0;
      t3 = 0;
    }//negro

    aux = int(random(7)); // **

    if(aux==0) // transmision de t por el canal con ruido e obteniendo r (r = t+e) *** // 000
    {
      r1 = (t1+0)%2;
      r2 = (t2+0)%2;
      r3 = (t3+0)%2;
    }
    else if(aux==1) // 001
    {
      r1 = (t1+0)%2;
      r2 = (t2+0)%2;
      r3 = (t3+1)%2;
    }
    else if(aux==2) // 010
    {
      r1 = (t1+1)%2;
      r2 = (t2+1)%2;
      r3 = (t3+0)%2;
    }
    else if(aux==3) // 100
    {
      r1 = (t1+1)%2;
      r2 = (t2+0)%2;
      r3 = (t3+0)%2;
    }
    else if(aux==4) // 101
    {
      r1 = (t1+0)%2;
      r2 = (t2+1)%2;
```

```
    r3 = (t3+0)%2;
  }
  else if(aux==5) // 110
  {
    r1 = (t1+0)%2;
    r2 = (t2+0)%2;
    r3 = (t3+0)%2;
  }
  else if(aux==6) // 011
  {
    r1 = (t1+0)%2;
    r2 = (t2+0)%2;
    r3 = (t3+0)%2;
  }
  else if(aux==7) // 111
  {
    r1 = (t1+0)%2;
    r2 = (t2+0)%2;
    r3 = (t3+0)%2;
  }

  for(j=1; j>=0; j--) // decodificacion de r (decodificacion por mayoria de votos) ****
  {
    P = p1*p2*p3;
    if(r1==j)
    {
      p1 = 1-f;
    }
    else
    {
      p1 = f;
    }

    if(r2==j)
    {
      p2 = 1-f;
    }
    else
    {
      p2 = f;
    }

    if(r3==j)
    {
      p3 = 1-f;
    }
    else
    {
      p3 = f;
    }
  }

  R = P/(p1*p2*p3); // calculo de cociente de probabilidades *****
  println(R);

  if(R<1.0) // decision optima ******
  {
    b = 0;
  }//negro
  else if(R>1)
  {
```

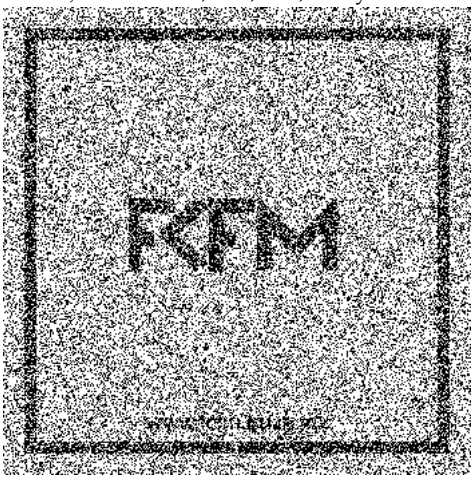
```
        b = 255;
      }//blanco

      img.pixels[i] = color(b,b,b); // bit salida
      R = 1;
      i++;
    }
    image(img,0,0);
    save("img_salida.png");
  }
```

a) Ruidos 001, 010 y 100.

b) Ruidos 001, 010, 100 y 110.

c) Ruidos 001, 010, 100, 110 y 111.



d) Ruidos 001, 010, 100, 110, 011 y 101.

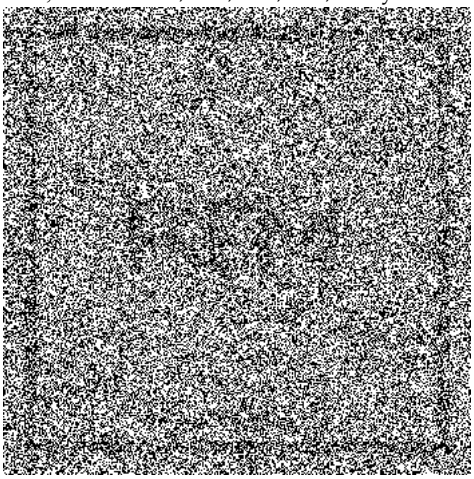


Tabla 5: Imagenes de salida con diferentes ruidos en el canal.

En la Tabla 5, mostramos diferentes salidas de la ejecución del programa, Código 1. Al observar la imagen a), vemos que los bloques de ruidos contenían un solo error, pero al pasar por la decodificación estos fueron corregidos y la imagen resulto ser la misma que la trasmitida, en cambio para las imágenes b), c) y d) no ocurrió lo mismo, ya que sabemos que está decodificación aplicada es capaz de corregir un solo error, al encontrar más, esta decodificación ya no es tan efectiva. Este proceso de codificación y decodificación estudiado, en sí nos ayuda un poco con el problema de la comunicación, sin embargo no es del todo confiable, por lo que es necesario estudiar otros algoritmos de codificación y decodificación.

1.3 CÓDIGOS DE BLOQUE: CÓDIGOS DE HAMMING

Nos gustaría comunicarnos con probabilidad de error muy pequeña y con una tasa grande. ¿Podemos mejorar los códigos de repetición? ¿Qué ocurre si añadimos redundancia a los bloques de datos en lugar de codificar un bit a la vez? Ahora estudiaremos un código de bloque simple que sigue siendo un código detector-corrector de error.

Un **código de bloque** es una regla para convertir una secuencia de bits fuente s de longitud k, en una secuencia transmitida t de N bits de longitud. A esta secuencia transmitida t se le llamará **palabra-código**. Para añadir redundancia, hacemos N más grande que k. En un código de bloque lineal, los bits extras $M = N - k$ están en función de los k bits originales; estos bits extras son conocidos como **bits de chequeo de paridad**. Además, la razón o cociente de proporcionalidad $R = \frac{k}{N}$, es llamada la **tasa** del código.

Un ejemplo de un código de bloque lineal es el $[7,4]$-*código de Hamming*, el cual transmite $N = 7$ bits por cada $k = 4$ bits fuentes, tiene $M = N - k = 7 - 4 = 3$ bits de chequeo de paridad, y una tasa de $R = \frac{k}{N} = \frac{4}{7}$.

Hablando un poco sobre la familia de códigos de Hamming, estos códigos fueron descubiertos por dos personas por Marcel Golay en 1949 y por Richard Hamming en 1950. Quizás esta familia sea la más famosa de todos los códigos detectores-correctores de errores. Además, son lineales, fáciles de codificar y decodificar. En particular, todos los códigos de Hamming binarios son equivalentes a los códigos cíclicos, y algunos códigos de Hamming no binarios también son equivalentes a los códigos cíclicos. En los siguientes apartados trabajaremos con el $[7,4]$-código de Hamming binario.

1.3.1 Codificación del $[7,4]$-código de Hamming

El $[7,4]$-código de Hamming es un código de bloque lineal, en el que por cada secuencia fuente s de longitud $k = 4$, es decir, s está constituido de 4 bits (recodar que seguimos trabajando con los dígitos binarios 0 y 1), se transmitirá una secuencia t (palabra-código) de longitud $N = 7$ bits, donde los $M = 3$ bits de chequeo de paridad estarán en función de los $k = 4$ bits fuentes. La operación de codificación para el $[7,4]$-código de Hamming está mostrado en la siguiente representación gráfica, Figura 5, [Mac03].

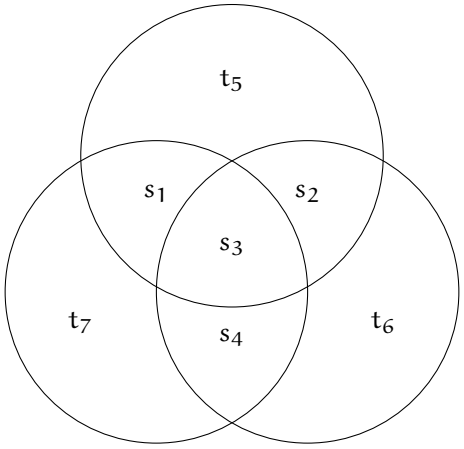


Figura 5: Representación gráfica de la codificación del $[7,4]$-código de Hamming.

Antes de continuar expliquemos que la **paridad** es la cualidad de ser par o impar de cualquier número, en el caso de una secuencia binaria tenemos:

- Paridad par: el valor asignado es 0 y esto se cumple cuando el número de unos que tengamos sea par o dicho de otra manera que la suma de esos unos sea 0 módulo 2.
- Paridad impar: el valor asignado es 1 y esto se cumple cuando el número de unos que tengamos sea impar o que la suma de esos unos sea 1 módulo 2.

El $[7,4]$-código de Hamming es descrito a continuación a partir de la Figura 5. Dado un arreglo $s = s_1s_2s_3s_4$ de cuatro bits, se obtendrá un arreglo $t = t_1t_2t_3t_4t_5t_6t_7$ de siete bits. El arreglo t de los siete bits transmitidos está representado en tres círculos intersectados. Los primeros cuatro bits transmitidos $t_1t_2t_3t_4$, son iguales a los cuatro bits fuentes $s_1s_2s_3s_4$, es decir, $t_1 = s_1$, $t_2 = s_2$, $t_3 = s_3$ y $t_4 = s_4$, cuyos bits son colocados en las intersecciones de los círculos como se muestra en la Figura 5. Luego, los tres bits de chequeo de paridad $t_5t_6t_7$ son colocados de tal manera que la suma de los bits dentro del círculo siempre de 0 usando aritmética módulo 2: el primer bit de chequeo de paridad t_5 es 0 sí la suma de los primeros tres bits es par, y 1 sí la suma es impar, claramente usando aritmética módulo 2, $s_1 + s_2 + s_3 = t_5$ (mód 2), entonces de acuerdo a nuestra representación gráfica el primer círculo en sentido de las manecillas del reloj representa el primer chequeo de paridad; el segundo bit de chequeo de paridad t_6 es la paridad de los tres últimos bits fuentes, $s_2 + s_3 + s_4 = t_6 \pmod 2$, con lo cual el segundo círculo representa al segundo chequeo de paridad; y el tercer bit de chequeo de paridad t_7 es la paridad de los bits fuentes uno, tres y cuatro, $s_1 + s_3 + s_4 = t_7 \pmod 2$, y el tercer círculo representa al tercer chequeo de paridad. Y así, obtenemos la palabra-código $t = s_1s_2s_3s_4t_5t_6t_7$, la cual es obtenida a partir de la codificación por el $[7,4]$-código de Hamming.

Ejemplo 1.6 Como un ejemplo, en la Figura 6 se muestra la palabra-código transmitida $t = 1000101$ que fue obtenida para el caso $s = 1000$.

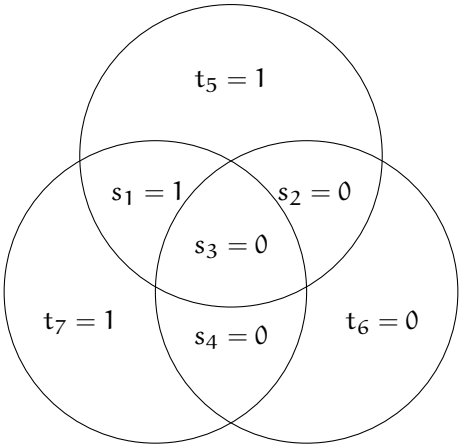


Figura 6: Palabra-código $t = 1000101$, secuencia fuente $s = 1000$.

Es fácil ver, que los primeros cuatro bits de t son los mismos bits de s, y luego los tres bits de chequeo de paridad 101 se obtienen de las respectivas paridades de los círculos, es decir, $s_1 + s_2 + s_3 = 1 + 0 + 0 = 1$ (mód 2) así $t_5 = 1$; $s_2 + s_3 + s_4 = 0 + 0 + 0 = 0$ (mód 2) luego $t_6 = 0$: y $s_1 + s_3 + s_4 = 1 + 0 + 0 = 1$ (mód 2) obteniendo $t_7 = 1$. Por eso la palabra-código obtenida de $s = 1000$ es $t = 1000101$.

La Tabla 6 muestra las palabras-código t generadas por cada una de las $2^4 = 16$ combinaciones de los cuatro bits fuente, las cuales son las secuencias fuente s. Esas palabras-código tienen la propiedad especial de que cualquier par difiere uno del otro en al menos tres bits, que posteriormente se definirá como distancia mínima.

s	t	s	t	s	t	s	t
0000	0000000	0100	0100110	1000	1000101	1100	1100011
0001	0001011	0101	0101101	1001	1001110	1101	1101000
0010	0010111	0110	0110001	1010	1010010	1110	1110100
0011	0011100	0111	0111010	1011	1011001	1111	1111111

Tabla 6: Las 16 palabras-código t del $[7,4]$-código de Hamming generadas por las secuencias fuente s.

Si denotamos a $\mathcal{C}$ como el conjunto de las palabras-código, entonces el conjunto de las palabras-código para el $[7,4]$-código de Hamming es:

$$\begin{aligned}\mathcal{C} = \{&0000000, 0001011, 0010111, 0011100, 0100110, 0101101, 0110001, 0111010,\\ &1000101, 1001110, 1010010, 1011001, 1100011, 1101000, 1110100, 1111111\}.\end{aligned} \tag{1.5}$$

1.3.2 Visión general de codificación para el [7,4]-código de Hamming

Debido a que el [7,4]-código de Hamming es un código lineal, podemos escribirlo compactamente en términos de matrices como sigue.

Dada la secuencia fuente $s \in \mathbb{F}_2^4$, la palabra-código $t \in \mathcal{C}$ donde $\mathcal{C}$ es el conjunto de las palabras-código es obtenida desde s a partir de una operación lineal, la cual iremos construyendo. Sea s un vector renglón de longitud $k = 4$:

$$s = (s_1, s_2, s_3, s_4).$$

Como describimos en la Figura 5, la palabra-código t de longitud $N = 7$, $t = (t_1, t_2, t_3, t_4, t_5, t_6, t_7)$ un vector renglón, está determinada por s, con respecto a las relaciones:

$$\begin{cases} t_1 = s_1 \\ t_2 = s_2 \\ t_3 = s_3 \\ t_4 = s_4 \\ t_5 = s_1 + s_2 + s_3 \\ t_6 = s_2 + s_3 + s_4 \\ t_7 = s_1 + s_3 + s_4 \end{cases}$$

ahora, consideremos a s_1, s_2, s_3 y s_4 variables libres tales que $s_1, s_2, s_3, s_4 \in \mathbb{F}_2 = \{0, 1\}$. Para lo siguiente usaremos el superíndice T que significa transpuesto(a). Entonces:

$$t^T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_1 + s_2 + s_3 \\ s_2 + s_3 + s_4 \\ s_1 + s_3 + s_4 \end{pmatrix} = s_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} + s_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Denotemos lo siguiente:

$$g_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \qquad g_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \qquad g_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{y} \qquad g_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}. \tag{1.6}$$

Los vectores g_1, g_2, g_3 y g_4 son linealmente independientes, y dichos vectores generan a $\mathcal{C}$ lo cual denotamos por $\mathcal{C} = [\{g_1, g_2, g_3, g_4\}]$, eso implica que $\mathcal{B} = \{g_1, g_2, g_3, g_4\}$ es base de $\mathcal{C}$ y así la $\dim \mathcal{C} = 4$. Luego, la matriz generadora del código está formada por los vectores (1.6):

$$G = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}. \tag{1.7}$$

Notar que las primeras cuatro columnas de G forman la matriz identidad de tamaño 4×4, esto es:

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.8}$$

y denotemos las columnas restantes como una matriz transpuesta P^T donde $P \in \mathcal{M}_{3\times 4}(\mathbb{F}_2)$ tal que:

$$P^T = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}. \tag{1.9}$$

Entonces, con lo analizado escribimos a la matriz generadora $G \in \mathcal{M}_{7\times 4}(\mathbb{F}_2)$ (1.7) como:

$$G = \begin{bmatrix} I_4 & | & P^T \end{bmatrix}, \qquad \text{con } I_4 \text{ definida en (1.8) y } P^T \text{ definida en (1.9).} \tag{1.10}$$

Definida la matriz G, podemos escribir lo siguiente:

$$sG = (s_1, s_2, s_3, s_4) \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_1 + s_2 + s_3 \\ s_2 + s_3 + s_4 \\ s_1 + s_3 + s_4 \end{pmatrix}^T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \end{pmatrix}^T = t.$$

Por lo tanto, todas las palabras-código t son obtenidas de la siguiente operación lineal:

$$t = sG, \qquad \text{con } s \in \mathbb{F}_2^4 \text{ y G la matriz generadora definida en (1.10),} \tag{1.11}$$

dicha operación de codificación (1.11) usa aritmética módulo 2, es decir, la suma y multiplicación definidas sobre el campo $\mathbb{F}_2 = \{0, 1\}$:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Más aún, en la operación de codificación (1.11) asumimos que s y t son vectores renglón, ahora, si s y t fuesen vectores columna, entonces esta operación es remplazada por:

$$t = G^T s, \tag{1.12}$$

donde G^T es la matriz generadora transpuesta del código, dada como:

$$G^T = \begin{bmatrix} I_4 \\ P \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Ambas relaciones (1.11) y (1.12) determinan el mismo conjunto de soluciones $\mathcal{C}$ de las palabras-código (1.5) para el $[7,4]$-código de Hamming. $\mathcal{C}$ constituye un subespacio del espacio vectorial $\mathbb{F}_2^7$ sobre el campo $\mathbb{F}_2$, llamado código de bloque lineal binario.

1.3.3 Decodificación del $[7,4]$-código de Hamming

Cuando se estudia una codificación más complicada $s \to t$, la tarea de decodificar el vector recibido r se vuelve menos simple. Debemos recordar que cualquiera de los bits pueden ser intercambiados, incluidos los bits de paridad. Si asumimos que el canal es un canal simétrico binario (CSB) y que todos los vectores fuentes son equiprobables, entonces la decodificación óptima identifica el vector fuente s cuya codificación $t(s)$ difiere del vector recibido r en pocos bits.

Se puede resolver el problema de decodificación, Figura 7, midiendo qué tan lejos está r de cada una de las dieciséis palabras-código en la Tabla 6 y entonces escogemos el más cercano. ¿Hay un camino más eficiente de encontrar el vector fuente más probable?

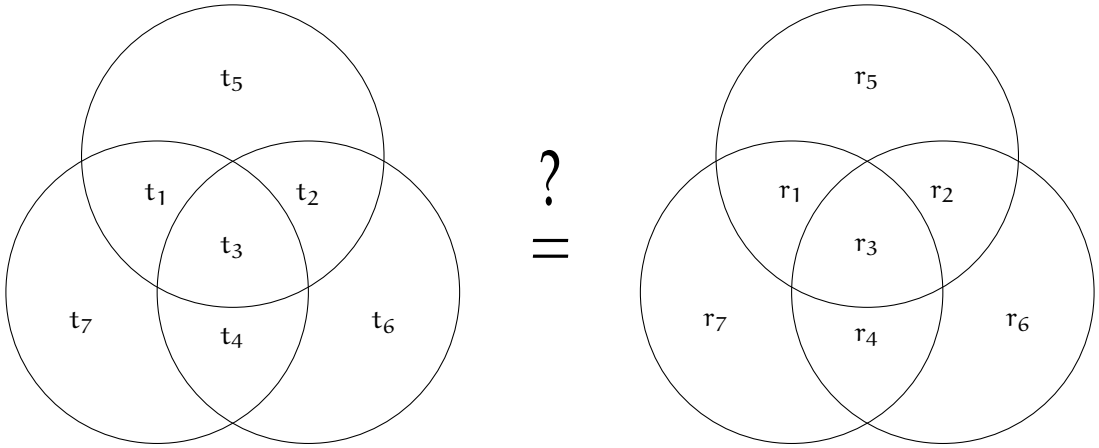


Figura 7: ¿t y r son iguales o qué tan similares son?

1.3.4 Síndrome de decodificación para el [7,4]-código de Hamming

Para el [7,4]-código de Hamming existe una solución gráfica al problema de decodificación, basado en la Figura 5 de codificación, [Maco3]. En las secciones anteriores 1.3.1 y 1.3.2 acabamos de estudiar como realizar la codificación para este código, obteniendo así la palabra-código $t = t_1t_2t_3t_4t_5t_6t_7$, ahora estudiaremos su decodificación, entonces dada la palabra-código t la cual es transmitida sobre un canal simétrico binario y es afectada por la presencia de ruido $e = e_1e_2e_3e_4e_5e_6e_7$, obteniendo así la secuencia recibida $r = r_1r_2r_3r_4r_5r_6r_7$, con $r = t + e \pmod 2$, esta secuencia r va ser decodificada para hallar el vector de salida $\hat{s} = \hat{s}_1\hat{s}_2\hat{s}_3\hat{s}_4$, y el vector de salida $\hat{s}$ tiene que ser lo más idéntico posible al vector fuente s.

La tarea de decodificación es encontrar los bits que fueron intercambiados, es decir, aquellos que no satisfacen las reglas de paridad. El patrón de error cuando no se satisfacen los chequeos de paridad es llamado **síndrome**, y pueden ser escritos como un vector binario $z = z_1z_2z_3$, cuyas componentes del vector tienen valor de, 1 sí el chequeo de paridad no es satisfecho ó 0 sí es satisfecho, pasando a través de los chequeos en el orden de los bits r_5, r_6 y r_7.

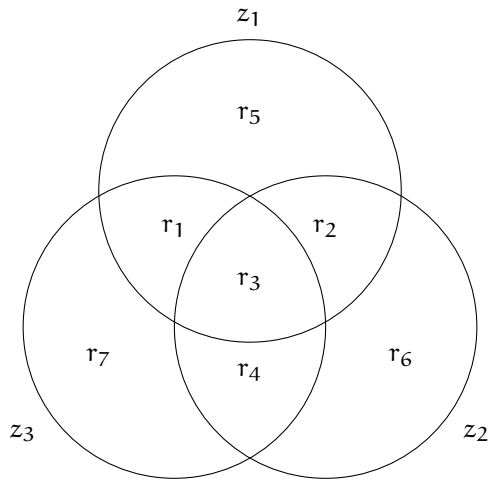


Figura 8: Vector recibido $r = r_1r_2r_3r_4r_5r_6r_7$ con su síndrome $z = z_1z_2z_3$.

La decodificación del [7,4]-código de Hamming es descrito a continuación con ayuda de la Figura 8. Nuestro objetivo es comprobrar que se cumplan las siguientes igualdades $r_1 + r_2 + r_3 = r_5$, $r_2 + r_3 + r_4 = r_6$ y $r_1 + r_3 + r_4 = r_7$, para lo cual basta con checar la paridad de $r_1 + r_2 + r_3 + r_5$, $r_2 + r_3 + r_4 + r_6$ y $r_1 + r_3 + r_4 + r_7$; y luego tomar una decisión óptima. Entonces al obtener la palabra recibida $r = r_1r_2r_3r_4r_5r_6r_7$ nos concentramos en los 4 bits que están en el primer círculo (en

sentido de las manecillas del reloj) de la Figura 8, es decir, en $r_1r_2r_3r_5$, entonces al sumar los bits tenemos que $r_1+r_2+r_3+r_5=0$ sí la cantidad de $1's$ es par, o bien, $r_1+r_2+r_3+r_5=1$ sí la cantidad de $1's$ es impar. Al obtener 0 de la suma diremos que la igualdad $r_1+r_2+r_3=r_5$ se cumple, pero en caso de obtener 1 de la suma diremos que la igualdad $r_1+r_2+r_3=r_5$ no se cumple y además que alguno de los bits involucrados posiblemente fue intercambiado. Al final de dicho análisis obtenemos el primer bit z_1 que forma parte del síndrome z. A esto nos referimos al decir que estamos realizando el chequeo de paridad. Ahora, este proceso se repite para los 4 bits del segundo círculo $r_2r_3r_4r_6$ con lo cual obtenemos el bit z_2, y para los 4 bits del tercer círculo $r_1r_3r_4r_7$ obteniendo el bit z_3. Una vez realizado esto tenemos al síndrome $z=z_1z_2z_3$ que nos dará una idea de cual es el bit intercambiado y poder tomar la decisión óptima.

Síndrome z	000	001	010	011	100	101	110	111
Bit intercambiado	ninguno	r_7	r_6	r_4	r_5	r_1	r_2	r_3

Tabla 7: Síndrome de decodificación para llevar a cabo la toma de decisión óptima.

La Tabla 7 muestra los resultados del algoritmo de decodificación para el $[7,4]$-código de Hamming, enlistando el vector síndrome z junto con el bit intercambiado, y obteniendo esto, es posible tomar la decisión óptima de decodificación, siempre y cuando solo se haya afectado a un bit del bloque.

Ejemplo 1.7 Como un primer ejemplo para ilustrar la decodificación del $[7,4]$-código de Hamming, asumimos que la palabra-código transmitida fue $t=1000101$, que es tomada del Ejemplo 1.6 de codificación, y que el ruido $e=0100000$ afecta al segundo bit, así el vector recibido es $r=t+e=1000101+0100000=1100101$. Escribimos el vector recibido $r=1100101$ dentro de los tres círculos como mostramos en la Figura 9.

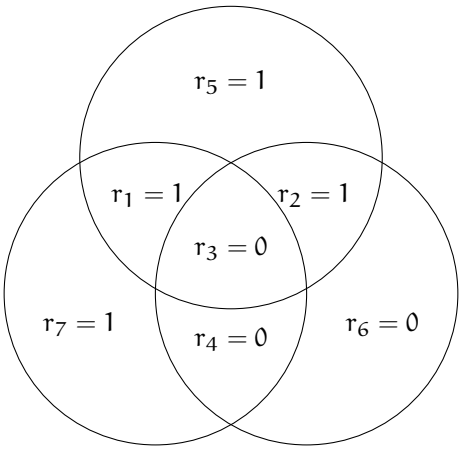


Figura 9: $r=1100101$ es el vector recibido.

Verifiquemos si el chequeo de paridad de cada círculo es par, en caso de no obtener chequeo de paridad par, los círculos de los cuales su paridad sea impar son trazados en líneas punteadas como se muestra en la Figura 10. En sí, el síndrome para cada chequeo de paridad de los círculos es $z=(1,1,0)$, porque los primeros dos círculos (en sentido de las manecillas del reloj) tienen chequeo de paridad impar, esto es, paridad 1, recordar que para calcular el chequeo de paridad del primer círculo debemos sumar los bits dentro de éste, así $r_1+r_2+r_3+r_5=1+1+0+1=1$, la suma de los bits del segundo círculo es $r_2+r_3+r_4+r_5=1+0+0+0=1$; y el tercer círculo tiene chequeo de paridad par, es decir, paridad 0, porque la suma de los bits de este círculo es $r_1+r_3+r_4+r_7=1+0+0+1=0$, (no olvidar que las operaciones usan aritmética módulo 2).

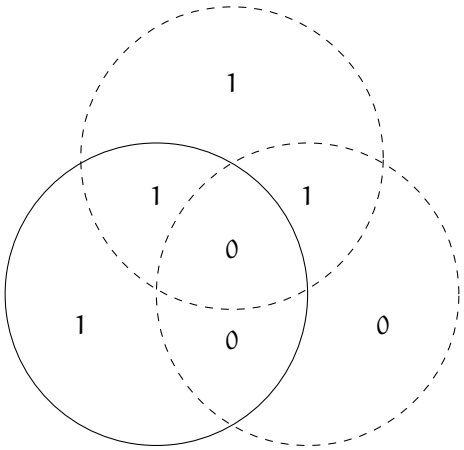


Figura 10: El síndrome es $z = (1, 1, 0)$, porque tenemos dos círculos de chequeo de paridad impar y solo uno de chequeo de paridad par.

Para resolver la tarea de decodificación, nos planteamos la siguiente pregunta: ¿podemos encontrar al bit que fue intercambiado dentro de todos los círculos de chequeo de paridad impar y fuera de todos los círculos de chequeo de paridad par?

Si es así, el bit intercambiado sería la causa o razón del síndrome observado. En este caso el bit $r_2 = 1$ afecta dentro de los dos círculos de chequeo de paridad impar y no afecta el círculo de chequeo de paridad par, Figura 11, ningún otro bit tiene esta propiedad, así $r_2 = 1$ es el único bit capaz de explicar el síndrome $z = (1, 1, 0)$.

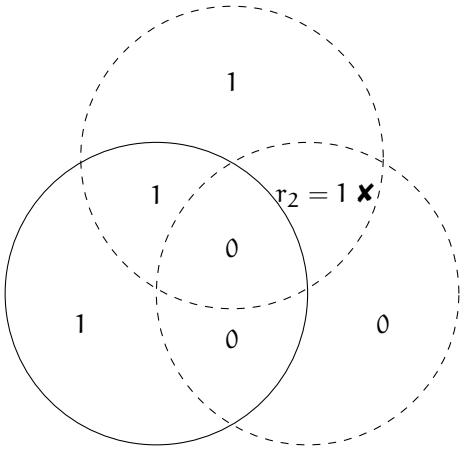


Figura 11: $r_2 = 1$ es el bit capaz de explicar el síndorme $z = (1, 1, 0)$.

Para hallar el bit que fue intercambiado, $r_2 = 1$ mostrado en la Figura 11, dada la Figura 9, procedemos de la siguiente manera: lo primero que debemos realizar (como ya lo hicimos) es calcular el chequeo de paridad de cada círculo, y obtuvimos que los dos primeros círculos tienen chequeo de paridad impar y el tercero chequeo de paridad par, esto es, síndrome $z = (1, 1, 0)$, Figura 10; ahora descartaremos todos los bits que están en el círculo de chequeo de paridad par, para este caso los bits descartados son $r_1 = 1$, $r_3 = 0$, $r_4 = 0$ y $r_7 = 1$, con esto los únicos sospechosos son $r_2 = 1$, $r_5 = 1$ y $r_6 = 0$ que están en los círculos de chequeo de paridad impar excluyendo la intersección con el círculo de chequeo de paridad par, ilustrado en la Figura 12. Finalmente teniendo los bits sospechosos, pondremos nuestra atención en el bit que está en la intersección, $r_2 = 1$, pues éste pareciera ser el bit que provoca que la paridad de estos círculos resulte ser impar, nuestra idea se basa en esto debido a que al haber realizado la codificación teníamos en cuenta que se obtenía un chequeo de paridad par en cada círculo. Por lo tanto, habremos encontrado el bit intercambiado, $r_2 = 1$, Figura 11.

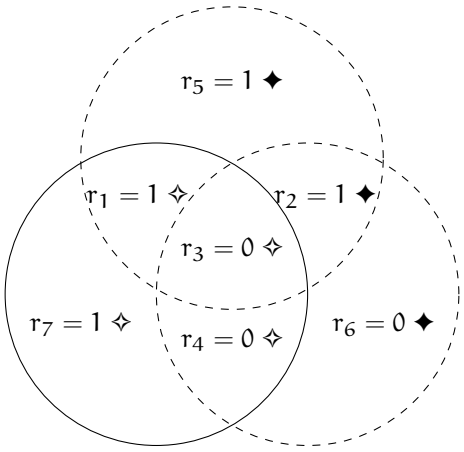


Figura 12: r_1, r_3, r_4 y r_7 son los bits descartados (✧); y r_2, r_5 y r_6 son los bits sospechosos (✦).

Y por último cambiaremos el valor de este bit $r_2 = 1$ por $r_2 = 0$, como en la Figura 13, obteniendo así, chequeo de paridad par en cada círculo.

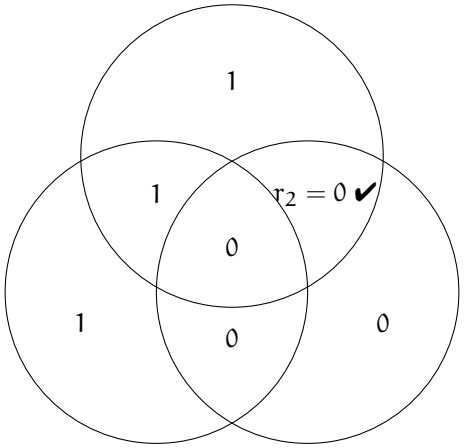


Figura 13: El valor de $r_2 = 1$ es cambiado por $r_2 = 0$, obteniendo así paridad par en cada círculo.

Entonces la secuencia resultante de la decodificación de $r = 1100101$ es 1000101, además, después de haber tomado la decisión óptima de decodificación concluimos que la secuencia de salida es $\hat{s} = 1000$, y es exactamente la misma que la secuencia fuente $s = 1000$.

Trabajemos con un par de ejemplos más.

Ejemplo 1.8 Tomando la misma secuencia transmitida del Ejemplo 1.7, $t = 1000101$, sea ahora la secuencia recibida $r = 1000001$, como se muestra en la Figura 14, aquí ocurre que el bit transmitido t_5 el cual es uno los bits de chequeo de paridad, es intercambiado por el ruido, el cual fue $e = 0000100$.

Realizando un análisis similar al anterior, vemos que $z = (1, 0, 0)$ pues el primer círculo tiene chequeo de paridad impar, y que el segundo y tercer círculo tienen chequeo de paridad par; descartando los bits que están en los círculos de chequeo de paridad par, solo nos queda el bit $r_5 = 0$, que es el posible bit sospechoso que afecta dentro de ese círculo de chequeo de paridad impar y fuera de los otros dos círculos de chequeo de paridad par, así r_5 es identificado como el único bit capaz de explicar el síndrome $z = (1, 0, 0)$.

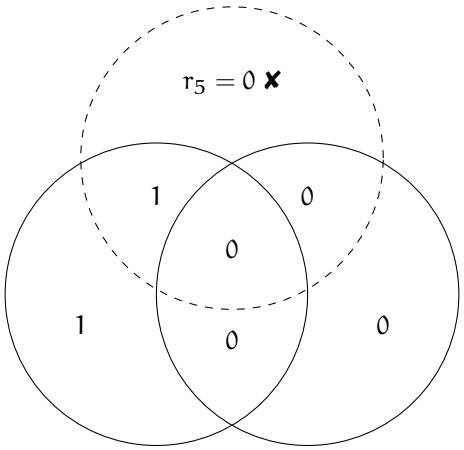


Figura 14: Decodificación de la secuencia recibida $r = 1000001$.

Cambiando este bit $r_5 = 0$ por $r_5 = 1$, obtenemos la secuencia resultante 1000101 de la decodificación de $r = 1000001$, Figura 15, entonces la secuencia de salida es $\hat{s} = 1000$, donde la decodificación volvió a ser óptima pues se corrigió el error, debido a que podemos verificar que la secuencia fuente $s = 1000$ es idéntica a la de salida.

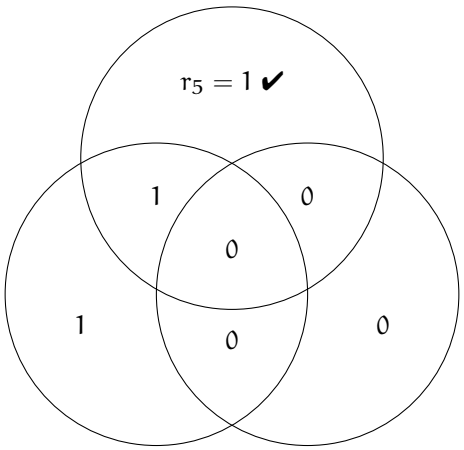


Figura 15: Resultado de la decodificación.

Ejemplo 1.9 Nuevamente tomemos a la secuencia transmitida $t = 1000101$, y ahora supongamos que el bit central r_3 es el bit intercambiado, es decir, es el afectado por el ruido $e = 0010000$, luego la secuencia recibida es $r = 1010101$. Continuando con su decodificación calculamos el chequeo de paridad de cada círculo, tenemos que los tres círculos tienen chequeo de paridad impar, como mostramos en la Figura 16. Ahora, para identificar al bit intercambiado nos fijaremos en el bit que está en la intersección de los tres círculos, ya que ese bit interviene en los tres chequeos de paridad, así el bit $r_3 = 1$ es identificado como el bit sospechoso que afecta dentro de los tres círculos y es capaz de explicar el síndrome $z = (1, 1, 1)$.

Una vez identificado el bit intercambiado $r_3 = 1$, modificaremos su valor por $r_3 = 0$, con lo que después de la decodificación de $r = 1010101$ obtenemos la secuencia resultante 1000101, Figura 17, entonces la secuencia de salida es $\hat{s} = 1000$ la misma que la secuencia fuente $s = 1000$, con lo que podemos decir que la decodificación volvió a resultar óptima.

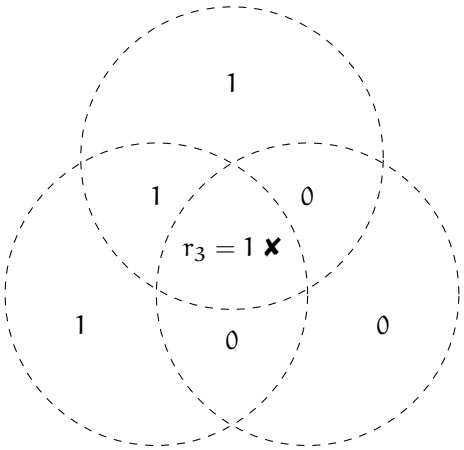


Figura 16: Cuando el bit r_3 es afectado por el ruido, ocurre que los círculos tienen paridad impar.

Finalmente, en la Figura 17 mostramos que la paridad de cada círculo vuelve a ser par pues en este caso el error fue una vez más corregido.

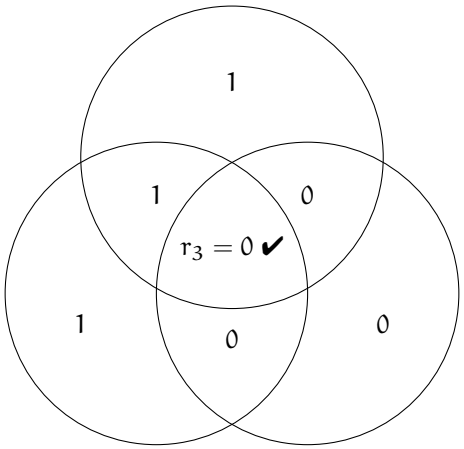


Figura 17: Al corregir el error la paridad de cada círculo vuelve a ser par.

Notemos que en los tres ejemplos anteriores, era un solo bit el intercambiado o afectado por el ruido, entonces si tratamos de intercambiar alguno de los 7 bits, encontraremos que un síndrome diferente es obtenido en cada caso, siete síndromes no-cero, uno para cada bit. Hay otro síndrome, el síndrome cero que indica que ningún bit fue intercambiado. Así, si el canal es un canal simétrico binario con un nivel de ruido pequeño f, la decodificación óptima de no intercambiar más de un bit, depende del síndrome, como mostramos en la Tabla 7. Sin embargo, también ocurre que cada síndrome podría ser causado por otro patrón de ruido, pero cualquier otro patrón de ruido que tenga el mismo síndrome podría ser menos probable porque eso involucraría un gran número de eventos ruidosos.

Nuestra siguiente pregunta sería: ¿qué ocurre si el ruido realmente intercambia más de un bit? Abordemos un ejemplo.

Ejemplo 2.0 Consideremos la situación cuando dos bits r_3 y r_7 son intercambiados, dado una vez más $t = 1000101$ con $s = 1000$, y sea el ruido $e = 0010001$ que afectará a dos bits, r_3 y r_7, tenemos que $r = 1010100$ como se ilustra en la Figura 18.

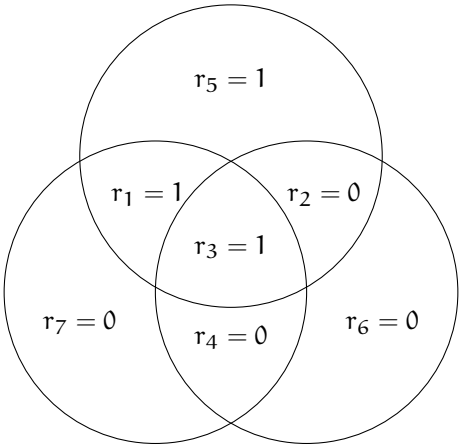


Figura 18: Situación cuando dos bits son afectados por el ruido, r_3 y r_7.

Calculando el síndrome $z = (1, 1, 0)$ obtenemos dos círculos de chequeo de paridad impar y uno de chequeo de paridad par, como se ilustra en la Figura 19.

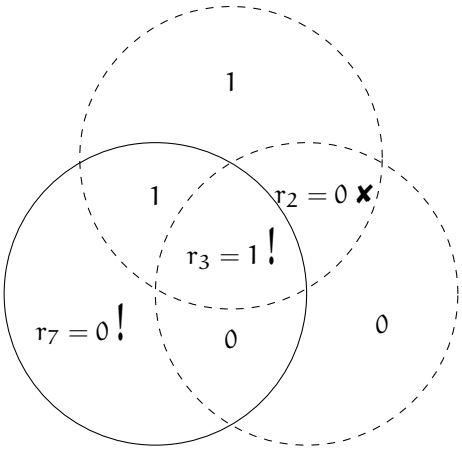


Figura 19: El síndrome para este caso es $z = (1, 1, 0)$ con lo cual se estaría diciendo que el bit sospechoso es r_2, pero sabemos que ese bit no es el intercambiado.

Como el síndrome es $z = (1, 1, 0)$, hace sospechoso al bit r_2 de acuerdo a la Tabla 7, entonces al aplicar el algoritmo de decodificación óptima para ese supuesto bit intercambiado, en realidad obtenemos como resultado una secuencia decodificada con tres errores como mostramos en la Figura 20. Estaríamos diciendo que la secuencia resultante de la decodificación de $r = 1010100$ es 1110100 cuando la secuencia transmitida fue $t = 1000101$, más aún la secuencia de salida resulta ser $\hat{s} = 1110$, pero en realidad sabemos que la secuencia fuente fue $s = 1000$, y en este caso concluimos que no resultó óptima la decodificación pues en lugar de corregir los dos errores, produjimos tres errores y obtuvimos algo muy distinto de la secuencia inicial.

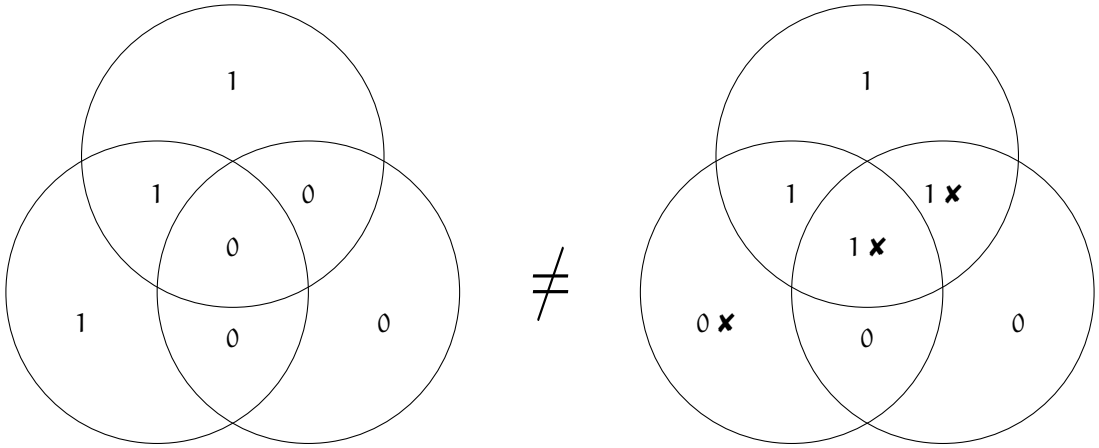


Figura 20: Decodificación no satisfactoria.

Con esto podemos decir, que si usamos el algoritmo de decodificación oṕtima, cualquier patrón de error con 2 bits intercambiados por el ruido, conducirá a una decodificación con un vector de 7 bits que contendrá tres errores.

1.3.5 Visión general de decodificación para códigos lineales: Síndrome de decodificación

También podemos describir el problema de decodificación para un código lineal en términos de matrices; nuevamente usaremos la notación del superíndice T que indica transpuesta(o) según sea el caso.

Dado r el vector recibido:

$$r = (r_1, r_2, r_3, r_4, r_5, r_6, r_7).$$

Los primeros cuatro bits recibidos, $r_1 r_2 r_3 r_4$, se proponen ser los cuatro bits fuente, y los tres últimos bits recibidos $r_5 r_6 r_7$, se proponen ser las paridades o los bits chequeos de paridad de los bits fuente, como está definido por la matriz generadora G (1.7). Evaluamos los tres bits de chequeo de paridad para los bits recibidos, $r_1 r_2 r_3 r_4$, y veamos si ellos corresponden a los tres bits recibidos $r_5 r_6 r_7$, es decir, tenemos que verificar si se cumplen las siguientes igualdades:

$$\begin{cases} r_5 = r_1 + r_2 + r_3 \\ r_6 = r_2 + r_3 + r_4 \\ r_7 = r_1 + r_3 + r_4 \end{cases}$$

El síndrome $z = (z_1, z_2, z_3)^T$ del vector recibido r, está dado por:

$$\begin{cases} z_1 = r_1 + r_2 + r_3 - r_5 \\ z_2 = r_2 + r_3 + r_4 - r_6 \\ z_3 = r_1 + r_3 + r_4 - r_7 \end{cases} \tag{1.13}$$

luego, por aritmética módulo 2:

$$\begin{aligned} r_1 + r_2 + r_3 + r_5 &= z_1 \\ r_2 + r_3 + r_4 + r_6 &= z_2 \\ r_1 + r_3 + r_4 + r_7 &= z_3 \end{aligned}$$

Del sistema de ecuaciones anterior (1.13), podemos obtener su representación matricial:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}. \tag{1.14}$$

Definimos a H como la matriz de chequeo de paridad del código dada por:

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}. \tag{1.15}$$

Notar que las tres últimas columnas de H forman la matriz identidad de tamaño 3×3:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.16}$$

y definamos las primeras cuatro columnas de H como la matriz P donde $P \in \mathcal{M}_{3\times 4}(\mathbb{F}_2)$:

$$P = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \tag{1.17}$$

luego, podemos escribir a la matriz de chequeo de paridad $H \in \mathcal{M}_{3\times 7}$ (1.15) como:

$$H = \begin{bmatrix} P & | & I_3 \end{bmatrix}, \qquad \text{con } P \text{ definida en (1.17) e } I_3 \text{ definida en (1.16).} \tag{1.18}$$

Antes de continuar con el análisis de la decodificación, es importante observar que la matriz generadora G del código, está asociada con la matriz de chequeo de paridad H del código. Recordar que G fue definida en (1.10) como:

$$G = \begin{bmatrix} I_4 & | & P^T \end{bmatrix},$$

donde I_4 es la matriz identidad 4×4 y $P^T = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ con $P \in \mathcal{M}_{3\times 4}(\mathbb{F}_2)$. Ahora, la matriz transpuesta de P^T es $P = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$, y es la misma P definida en (1.17), la cual forma parte de la matriz de chequeo de paridad $H = \begin{bmatrix} P & | & I_3 \end{bmatrix}$ definida en (1.18).

Esto es, si tenemos $G = \begin{bmatrix} I_4 & | & P^T \end{bmatrix}$ donde I_4 es la matriz identidad de tamaño 4×4 y $P \in \mathcal{M}_{3\times 4}(\mathbb{F}_2)$ entonces $H = \begin{bmatrix} P & | & I_3 \end{bmatrix}$ donde $P \in \mathcal{M}_{3\times 4}(\mathbb{F}_2)$ e I_3 es la matriz identidad de tamaño 3×3, y viceversa. Cuando esto ocurre, para un código lineal binario, se dice que G es la matriz generadora estándar del código y H es la matriz de chequeo de paridad estándar del código.

Continuando con el análisis de la decodificación, definida la matriz H, podemos escribir la representación matricial (1.14) como:

$$Hr^T = z,$$

así, el cálculo del vector síndrome z es una operación lineal.

- Si el síndrome es cero, es decir, $z = (0,0,0) = \vec{0}$, lo cual ocurre si los tres bits de chequeos de paridad son satisfechos, esto es:

  $$Hr^T = \vec{0},$$

  entonces el vector recibido r es una palabra-código en $\mathcal{C}$, y la decodificación más probable está dada al leer esos primeros cuatro bits, esto es, el vector de salida sería $\hat{s} = (r_1, r_2, r_3, r_4)$.

- Si el síndrome no es cero, es decir, $z \neq \vec{0}$, lo cual ocurre si alguno de los tres bits de chequeos de paridad no fue satisfecho, esto es:

  $$Hr^T \neq \vec{0},$$

  entonces las secuencias de ruido para este bloque no fue cero, y el síndrome es el punto más probable de patrón de error.

Podemos decir que todas las palabras-código $t = sG$ del código $\mathcal{C}$ satisfacen:

$$Ht^T = \vec{0}. \tag{1.19}$$

Luego, desde que el vector recibido r está dado por $r = t + e = sG + e$, donde e es el ruido, y para el problema de la decodificación del síndrome tenemos que verificar:

$$\begin{aligned} Hr^T = z &\Rightarrow H(t+e)^T = z \quad \text{(pues } r = t + e\text{)} \\ &\Rightarrow Ht^T + He^T = z \quad \text{(usando la propiedad de linealidad)} \\ &\Rightarrow \vec{0} + He^T = z \quad \text{(por (1.19))} \\ &\Rightarrow He^T = z. \end{aligned}$$

Con lo que el problema de decodificación se reduce a encontrar el vector de ruido más probable e que satisface la ecuación:

$$He^T = z.$$

El algoritmo de decodificación que resuelve este problema es llamado un **decodificador de máxima probabilidad**.

1.3.6 ¿Qué ocurre con la probabilidad de error para el [7,4]-código de Hamming?

Resulta ser que este código detector-corrector de error, nos referímos al [7,4]-código de Hamming, al igual que el código de repetición $\mathcal{R}_3$, sólo es capaz de corregir un error, entonces posiblemente cada vector recibido de longitud de 7 bits es alguna de las palabras-código de $\mathcal{C}$, u ocurre que algunos de sus bits fueron intercambiados y no fue exitosa la decodificación, obteniendo un vector de salida incorrecto. Las probabilidades de error no cambian mucho, pero las tasas de los códigos sí difieren más notablemente pues para el código de repetición $\mathcal{R}_3$ la tasa es $R = \frac{k}{N} = \frac{1}{3}$ y para el [7,4]-código de Hamming es $R = \frac{k}{N} = \frac{4}{7}$, con lo que somos capaces de decir que sí se consigue mejorar los códigos de repetición usando los códigos de bloque.

Como ya hemos estudiado hay tres restricciones de paridad, y cada una podría o no ser satisfecha, luego se tienen $2^3 = 8$ síndromes distintos. Pueden ser divididos en siete síndromes no cero, uno para cada bit de los patrones de error, y el síndrome cero, correspondiente al caso en el que el ruido fue cero y ningún bit fue afectado.

La tarea de decodificación óptima no actúa si el síndrome es cero, de suceder lo contrario usamos los resultados del algoritmo de decodificación mostrado en la Tabla 7 para hallar el bit sospechoso que fue intercambiado.

Obtenemos un error de decodificación si el vector de salida $\hat{s}$ no es igual al vector fuente s, es decir, si uno o más de los cuatro bits de salida $\hat{s}_1$, $\hat{s}_2$, $\hat{s}_3$, $\hat{s}_4$ decodificados no corresponden a los bits fuente s_1, s_2, s_3, s_4. Luego, la probabilidad de error por bloque P_B es la probabilidad de que uno o más bits decodificados en un bloque falle al hacer la correspondencia con los bits fuente:

$$P_B = P(\hat{s} \neq s).$$

La probabilidad de error del bit P_b es la probabilidad promedio de que un bit decodificado falle al hacer la correspondiencia al bit fuente:

$$P_b = \frac{1}{k}\sum_{i=1}^{k} P(\hat{s}_i \neq s_i).$$

En el caso del $[7,4]$-código de Hamming, un error de decodificación ocurrirá cuando el ruido ha cambiado más de un bit en un bloque de 7. La probabilidad de error por bloque es por tanto la probabilidad de que dos o más bits sean intercambiados en el bloque. Esa probabilidad es de orden $O(f^2)$, como ocurrió en la probabilidad de error para el código de repetición $\mathcal{R}_3$. Pero la buena noticia es que el código de Hamming se comunica con una mayor tasa, $R = \frac{4}{7}$, con lo cual mejoramos la transmisión del mensaje.

1.3.7 Simetría del $[7,4]$-código de Hamming

Para probar que el $[7,4]$-código de Hamming protege todos los bits igualmente, comencemos con la matriz de chequeo de paridad H_1:

$$H_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{1.20}$$

La simetría entre los siete bits transmitidos será más fácil de ver si reordenamos los siete bits usando la siguiente permutación σ:

$$\sigma = \begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 \\ t_5 & t_2 & t_3 & t_4 & t_1 & t_6 & t_7 \end{pmatrix} \tag{1.21}$$

Entonces reescribimos H_1 (1.20) como:

$$H_2 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \tag{1.22}$$

Al comparar las matrices H_1 en (1.20) y H_2 en (1.22), observamos que las matrices no son iguales, entonces es lógico pensar que el conjunto de soluciones para H_1 es distinto del conjunto de soluciones para H_2, y estaríamos obteniendo dos códigos $\mathcal{C}_1$ y $\mathcal{C}_2$, respectivamente. Y esto no sería lo más apropiado porque ya no trabajaríamos sobre el $[7,4]$-código de Hamming. Sin embargo, resulta ser que dos códigos $\mathcal{C}_1$ y $\mathcal{C}_2$ son *equivalentes* si uno se obtiene a partir del otro por una permutación de los componentes de sus palabras-código. De esta manera, H_1 y H_2 generan el mismo código (salvo equivalencia).

Ahora, si tomamos cualesquiera dos restricciones de paridad que satisfacen t de la nueva matriz H_2, (1.22), y los sumamos, conseguimos otra restricción de paridad. Por ejemplo, el renglón 1 de (1.22), $t_5 + t_2 + t_3 + t_1 = 0$ y el renglón 2 de (1.22), $t_2 + t_3 + t_4 + t_6 = 0$, y la suma de esas dos restricciones es:

$$t_5 + 2t_2 + 2t_3 + t_1 + t_4 + t_6 = 0 \tag{1.23}$$

usando aritmética módulo 2 simplificamos la ecuación anterior (1.23) obteniendo:

$$t_5 + t_1 + t_4 + t_6 = 0$$

la cual podremos añadirla a la matriz chequeo de paridad H_2 (1.22) como un cuarto renglón, el conjunto de vectores que satisfacen $H_2 t^T = 0$ no cambia, pues esta nueva restricción es resultado de la suma de dos restricciones de H_2. Por lo tanto definimos:

$$H_2' = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{1.24}$$

Así, el cuarto renglón de H_2' en (1.24) es la suma módulo 2 del primer y segundo renglón de H_2. Notar que el segundo, tercer y cuarto renglón son todos desplazamientos cíclicos del renglón de arriba. Si después de haber añadido la cuarta restricción redundante, descartamos la primera restricción, obtenemos una nueva matriz de chequeo de paridad:

$$H_2'' = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{1.25}$$

la cual aún satisface $H_2'' t^T = 0$ para todas las palabras-código. Esto se da, porque el conjunto de soluciones para H_2 y H_2'' son equivalentes, pues notemos que todas las columnas de H_2 han sido intercambiadas a la derecha, y la columna del extremo derecho ha reaparecido a la izquierda, esto es, una permutación cíclica de las columnas, obteniendo así a H_2'' (1.25).

Esto establece la simetría entre los siete bits, ya que si iteramos el procedimiento anteriormente dicho cinco veces más, podemos hacer un total de siete matrices diferentes para el mismo código original, bajo el término de equivalencia, donde cada matriz asigna a cada bit un rol distinto.

Pudiéramos también construir la matriz de súper-redundancia de siete renglones de chequeo de paridad para el código, la cual es:

$$H_2''' = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{1.26}$$

Esta matriz H_2''' (1.26) es redundante en el sentido que, al reducir la matriz a su forma escalonada, tenemos que el número de renglones linealmente independientes, es 3, luego el *rango*$(H_2''') = 3$, y además, su espacio generado es un espacio de dimensión 3.

Observamos que esta matriz es una matriz cíclica. Cada fila es un corrimiento cíclico hacia la derecha de la fila superior. Y la dicha matriz nos da lugar a la siguiente definición.

1.4 CÓDIGOS CÍCLICOS

Si hay un ordenamiento de los bits de las palabras-código, tal que un código lineal tiene una matriz de chequeo de paridad cíclica, entonces el código es llamado un *código cíclico*.

Definición 1.2. *Un código lineal* $\mathcal{C} \subseteq \mathbb{F}_2^n$ *es un código cíclico si, para cada palabra-código* $c = (c_0, c_1, c_2, \ldots, c_{n-1}) \in \mathcal{C}$, *el corrimiento cíclico a la derecha de* c, *implica que* $(c_{n-1}, c_0, c_1, \ldots, c_{n-2}) \in \mathcal{C}$.

Por ejemplo, el siguiente código lineal con 16 palabras-código es un código cíclico, ya que satisface la Definición 1.2:

$$\begin{aligned} \mathcal{C}_1 = \{&0000000, 1110100, 0111010, 0011101, 1001110, 0100111, 1010011, 1101001, \\ &1011000, 0101100, 0010110, 0001011, 1000101, 1100010, 0110001, 1111111\}. \end{aligned}$$

Ahora, si usamos la permutación σ definida en (1.21), que indica que permutando la primer componente por la quinta componente de cada una de las palabras-código del código cíclico $\mathcal{C}_1$, obtenemos las palabras-código que se enlistan en la siguiente Tabla 8, y a tal código lo denotamos como $\mathcal{C}_2$:

$\mathcal{C}_1$	$\mathcal{C}_2$
0 0 0 0 0 0 0	0 0 0 0 0 0 0
1 1 1 0 1 0 0	1 1 1 0 1 0 0
0 1 1 1 0 1 0	0 1 1 1 0 1 0
0 0 1 1 1 0 1	1 0 1 1 0 0 1
1 0 0 1 1 1 0	1 0 0 1 1 1 0
0 1 0 0 1 1 1	1 1 0 0 0 1 1
1 0 1 0 0 1 1	0 0 1 0 1 1 1
1 1 0 1 0 0 1	0 1 0 1 1 0 1
1 0 1 1 0 0 0	0 0 1 1 1 0 0
0 1 0 1 1 0 0	1 1 0 1 0 0 0
0 0 1 0 1 1 0	1 0 1 0 0 1 0
0 0 0 1 0 1 1	0 0 0 1 0 1 1
1 0 0 0 1 0 1	1 0 0 0 1 0 1
1 1 0 0 0 1 0	0 1 0 0 1 1 0
0 1 1 0 0 0 1	0 1 1 0 0 0 1
1 1 1 1 1 1 1	1 1 1 1 1 1 1

Tabla 8: Permutación de las palabras-código del código cíclico $\mathcal{C}_1$ para obtener el código $\mathcal{C}_2$.

Al aplicar la permutación sobre $\mathcal{C}_1$, obtenemos 16 nuevas palabras código, pero en realidad este nuevo código lineal $\mathcal{C}_2$ es el $[7,4]$-código de Hamming, el cual podemos verificar con el conjunto dado en (1.5). Esto es, el $[7,4]$-código de Hamming, consiste de 7 desplazamientos cíclicos de las palabras-código 1110100 y 1011000, y las palabras-código 0000000 y 1111111. De tal manera, el $[7,4]$-código de Hamming es equivalente a un código cíclico bajo una permutación.

1.5 GRAFOS CORRESPONDIENTES A CÓDIGOS

Al diseñar un código detector-corrector de error y un algoritmo de decodificación, posiblemente nos demos cuenta que es más fácil inventar un nuevo código que hallar su decodificación óptima. Existen varios caminos para diseñar códigos, y lo que sigue es una de las maneras de construir un nuevo código.

En la subsección 1.3.1 introdujimos una representación gráfica del $[7,4]$-código de Hamming, Figura 5. Si observamos esa figura veremos que en cada uno de los círculos hay cuatro bits de la palabra-código transmitida, estos cuatro bits determinan el chequeo de paridad de cada uno de los círculos, entonces lo que haremos es reemplazar cada uno de los bits transmitidos por un *nodo bit* y conectar los cuatro *nodos bit* que están en cada círculo por un *nodo chequeo de paridad*, como en la Figura 21.

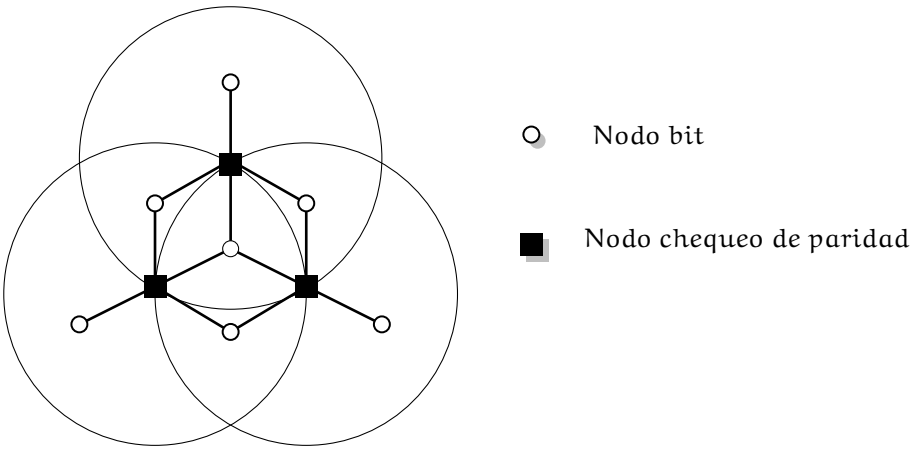


Figura 21: Cada círculo tiene un nodo chequeo de paridad que conecta a cuatro nodos bit.

Entonces, al prescindir de los tres círculos y nombrar a cada nodo bit y nodo chequeo de paridad, obtenemos una representación del $[7,4]$-código de Hamming por un grafo bipartito como se muestra en la Figura 22, [Mac03].

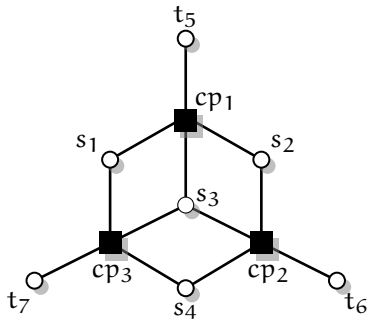


Figura 22: Grafo bipartito del $[7,4]$-código de Hamming.

Los 7 círculos pequeños nombrados **nodos bit** son los 7 bits transmitidos. Los 3 cuadrados pequeños son los **nodos chequeo de paridad**, es decir, los nodos restricción (no deben ser confundidos con los 3 bits de chequeo de paridad, los cuales son los 3 círculos más periféricos). El grafo es un grafo bipartito porque sus nodos se subdividen en 2 clases -bits y chequeos- y las aristas unen nodos de clases diferentes.

Muchos códigos pueden ser representados mediante un grafo bipartito. La conexión que existe entre los códigos lineales y los grafos bipartitos es la matriz de chequeo de paridad H, en donde cada nodo chequeo de paridad corresponde a una fila de H y cada nodo bit corresponde a una columna de H, y por cada 1 en H, existe una arista entre el correspondiente par de nodos. Entonces, dada la matriz de chequeo de paridad H definida en (1.15):

$$H = \begin{array}{c} \\ cp_1 \\ cp_2 \\ cp_3 \end{array} \begin{array}{c} \begin{array}{ccccccc} s_1 & s_2 & s_3 & s_4 & t_5 & t_6 & t_7 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

podemos ver cómo la matriz que describe al $[7,4]$-código de Hamming también puede describir al grafo bipartito ilustrado en la Figura 22.

Una vez establecida esta conexión entre códigos y grafos, un camino para construir códigos lineales es simplemente pensar en un grafo bipartito.

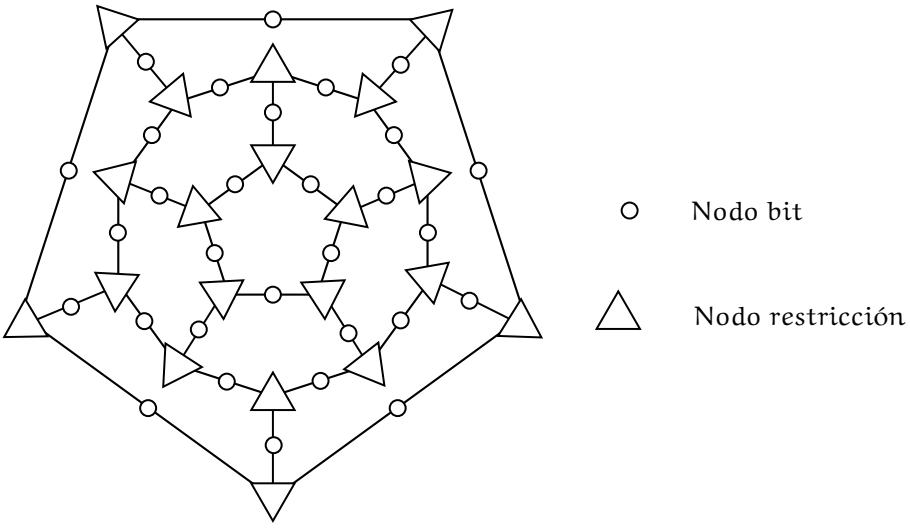


Figura 23: Grafo bipartito asociado al $[30,11]$-código dodecaedro.

Por ejemplo, un grafo bipartito interesante puede ser formado por polígonos regulares, como dos pentágonos y un decágono. Ahora, si por cada vértices colocamos un nodo restricción y a la mitad de cada arista un nodo bit, conseguimos el grafo bipartito ilustrado en la Figura 23.

Esta construcción define una matriz de chequeo de paridad en la cual cada columna tiene peso 2 y cada renglón tiene peso 3. El peso de un vector binario es el número de unos que éste contiene. Este código tiene $N = 30$ nodos bit, y al parecer $M = 20$ nodos restricción, pero solo hay $M = 19$ nodos restricción independientes; el vigésimo nodo restricción es redundante pues dado que 19 nodos restricción se satisfacen, entonces el vigésimo se cumple; luego el número de bits fuente es $k = N - M = 30 - 19 = 11$. Además, con el grafo bipartito podemos formar un poliedro con doce caras, un dodecaedro. Por lo tanto, el código que define se le conoce como $[30, 11]$-código dodecaedro.

Es díficil hallar un algoritmo de decodificación para este código, pero podemos estimar las probabilidades de error encontrando las palabras-código de peso más bajo. Si intercambiamos todos los bits rodando una cara del dodecaedro original, entonces todos los nodos restricción serán satisfechos; así el código tiene 12 palabras-código de peso 5, uno por cada cara. Como el peso más bajo de las palabras-código es 5, decimos que el código tiene distancia $d = 5$, el $[7, 4]$-código de Hamming tiene distancia $d = 3$ y corrige todos los errores de un solo bit intercambiado. Un código con distancia $d = 5$ es capaz de corregir todos los errores dobles de los bits intercambiados, pero hay algunos errores triples de bits intercambiados que no puede corregir.

Así, el error de probabilidad de este código, suponiendo un canal simétrico binario, es menor, al menos para bajos niveles de ruido f, por un término de orden f^3:

$$12 \binom{5}{3} f^3 (1-f)^{27}.$$

Algunos códigos lineales tienen una descripción gráfica simple, pero sus grafos son más complicados de analizar debido a que sus aristas se cruzan mucho, como el pequeño $[16, 4]$-código de Gallager cuyo grafo bipartito se ilustra en la Figura 24, dicho código de bloque lineal tiene longitud $N = 16$ nodos bit, $M = N - k = 16 - 4 = 12$ nodos restricción y tasa $R = \frac{k}{N} = \frac{4}{16} = \frac{1}{4}$.

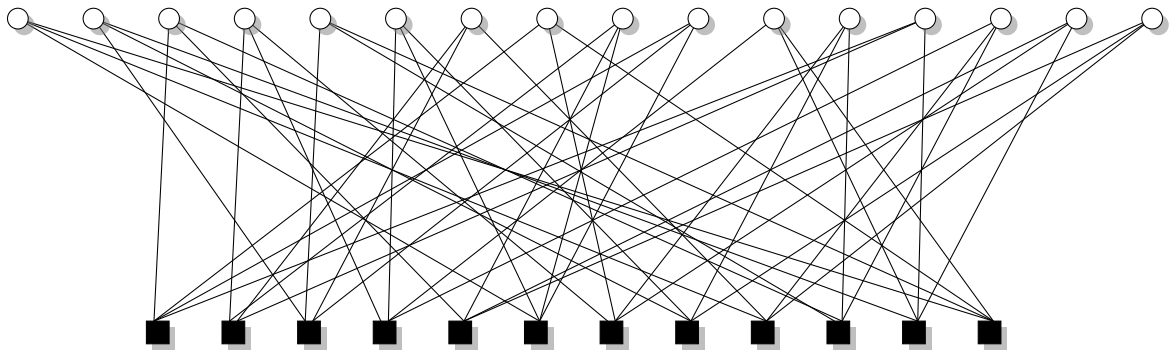


Figura 24: Grafo bipartito con tasa $R = \frac{1}{4}$ que representa al $[16, 4]$-código de Gallager (un código LDPC) de bloque de longitud $N = 16$ nodos bit y $M = 12$ nodos restricción.

Además, no hay razón para apegarse a los códigos lineales, ciertamente algunos *códigos no lineales*, códigos cuyas palabras-código no puede ser definidas por una ecuación lineal como $Ht^T = 0$, tienen muy buenas propiedades. Pero la codificación y decodificación de un código no lineal hace esas tareas aún más difíciles.

2 CÓDIGOS Y GRAFOS

El matemático, que se encuentra bajo su diluvio de símbolos, y trabaja, al parecer, con verdades puramente formales, puede aún alcanzar resultados de infinita importancia para nuestra descripción del universo físico.

– Karl Pearson –

En este capítulo introduciremos las nociones básicas de la teoría de códigos, la teoría de grafos, el concepto de un grafo expandido, enunciaremos el Teorema de codificación de Shannon, y la relación que existe entre los códigos lineales con los grafos bipartitos, dichos conceptos son esenciales para el estudio de los códigos LDPC.

2.1 ESTRUCTURA BÁSICA DE LOS CÓDIGOS

Las reglas bajo las cuales el codificador y el decodificador funcionan son especificadas por el código particular que es utilizado.

El tipo de código sobre el cual trabajaremos son códigos de bloque, estos códigos convierten una secuencia fuente u (mensaje fuente) de k dígitos de longitud, en una secuencia transmitida x (palabra-código) con n dígitos de longitud. Una vez que el mensaje fuente u es codificado en la palabra-código x, x es enviada por el canal, pero debido al ruido del canal x es afectada y el vector recibido es $y = x + e$, y quizá sea diferente de x. Con lo cual se define el vector error $e = y - x$ (patrón de error), el cual nos ayudará a determinar el síndrome de decodificación s para luego tomar la decisión óptima y así obtener la palabra de salida decodificada $\hat{u}$.

En esta sección se describe la estructura que tiene este tipo de códigos de bloque, para poder analizar su proceso de codificación y decodificación; los textos revisados para el desarrollo de dicha sección fueron [LA14], [McE04], [Rom92] y [RU08].

2.1.1 Conceptos básicos de un código

Comenzaremos con las definiciones y conceptos básicos de un código.

Definición 2.1. *Un código $\mathcal{C}$ de longitud n y cardinalidad finita M sobre un campo $\mathbb{F}$ es una colección de M elementos de $\mathbb{F}^n$, es decir,*

$$\mathcal{C} = \{x^{[1]}, \ldots, x^{[M]}\}, \quad x^{[m]} \in \mathbb{F}^n, \quad 1 \leqslant m \leqslant M, \tag{2.1}$$

así, $\mathcal{C}$ es un (n, M)-código.

Comentario 2.1. *Los elementos del código $\mathcal{C}$, $x^{[m]}$ son llamados palabras-código y el parámetro n es llamado la longitud de las palabras-código o simplemente la longitud del código.*

Ejemplo 2.1. *Sea el campo $\mathbb{F}_2 = \{0, 1\}$. El código de repetición binario de longitud $n = 3$ con $M = 2$ palabras-código es $\mathcal{C} = \{000, 111\}$.*

Comentario 2.2. *Es natural y conveniente usar campos finitos para realizar la codificación. Debido a esto, el campo sobre el cual trabajaremos es el campo binario $\mathbb{F}_2$, consistente de los elementos 0 y 1, con adición módulo 2 (2.2), y multiplicación módulo 2 (2.3).*

$+$	0	1
0	0	1
1	1	0

Suma sobre $\mathbb{F}_2 = \{0, 1\}$ (2.2)

·	0	1
0	0	0
1	0	1

Multiplicación sobre $\mathbb{F}_2 = \{0, 1\}$ (2.3)

En otras palabras, si usamos el campo $\mathbb{F}_2 = \{0, 1\}$ *entonces representamos información en términos de secuencias de bits (dígitos binarios).*

Ahora, demos la definición de código sobre $\mathbb{F}_2$.

Definición 2.2. *Sea* $\mathbb{F}_2^n$ *el conjunto de n-adas de elementos de* $\mathbb{F}_2$, $\mathbb{F}_2^n$ *es un espacio vectorial sobre el campo* $\mathbb{F}_2$. *Se dice que* $\mathcal{C}$ *es un código de longitud* n *sobre* $\mathbb{F}_2$ *si* $\mathcal{C}$ *es un subconjunto de* $\mathbb{F}_2^n$. *Un* (n, M)*-código* $\mathcal{C}$ *sobre* $\mathbb{F}_2$ *es un código de longitud* n *y tamaño* M.

Comentario 2.3. *A estos códigos sobre el campo* $\mathbb{F}_2$ *se les conoce como códigos binarios, y sus palabras-código como palabras-código binarias.*

Ejemplo 2.2. *Sea* $\mathbb{F}_2^7$ *el conjunto de todas las* n*-adas de longitud 7 sobre el campo binario* $\mathbb{F}_2 = \{0, 1\}$. *Y sea* $\mathcal{C} \subseteq \mathbb{F}_2^7$,

$$\begin{aligned}\mathcal{C} = \{&0000000, 0001011, 0010111, 0011100, 0100110, 0101101, 0110001, 0111010,\\ &1000101, 1001110, 1010010, 1011001, 1100011, 1101000, 1110100, 1111111\}.\end{aligned}$$

Decimos que $\mathcal{C}$ *es un código de longitud* $n = 7$ *y tamaño* $M = 16$ *palabras-código. Entonces,* $\mathcal{C}$ *es un* $(7, 16)$*-código binario.*

Ahora definiremos algunos parámetros importantes de los códigos.

Definición 2.3. *La tasa* R *de un código binario* $\mathcal{C}$ *se define como la proporción entre la longitud de los* k *bits del mensaje fuente* u *y la longitud de los* n *bits de la palabra-código* x. *Es decir, la tasa del código es*

$$R = \frac{k}{n} \tag{2.4}$$

que indica la capacidad que tiene el código $\mathcal{C}$ *en codificar* k *bits del mensaje fuente* u *en un total de* n *bits de la palabra-código* x.

Definición 2.4. *La distancia de Hamming entre dos vectores* $x, y \in \mathbb{F}_2^n$, *la cual denotamos por* $d(x, y)$, *se define como*

$$d(x, y) = |\{i : 1 \leqslant i \leqslant n, x_i \neq y_i\}|, \tag{2.5}$$

es decir, es el número de posiciones de las coordenadas en las cuales x *e* y *difieren.*

Ejemplo 2.3. *Sean* $x = (0, 1, 1, 1, 0, 1, 0)$, $y = (1, 0, 0, 1, 1, 1, 0)$ *vectores de* $\mathbb{F}_2^7$. *Luego, la distancia de Hamming entre* x *e* y *es*

$$d(x, y) = |\{1, 2, 3, 5\}| = 4,$$

ya que 1, 2, 3 *y* 5 *son las posiciones de las coordenadas en las que* x *e* y *difieren.*

Definición 2.5. *La distancia mínima de Hamming* $d(\mathcal{C})$ *de un código* $\mathcal{C}$, *se define como*

$$d(\mathcal{C}) = \text{mín}\{d(x, y) : x, y \in \mathcal{C}, x \neq y\}, \tag{2.6}$$

es decir, es el mínimo de todas las distancias para cada par de palabras-código.

Observación 2.1. *Para encontrar la distancia mínima* $d(\mathcal{C})$ *de un* (n, M)*-código* $\mathcal{C}$, *necesitamos calcular* $\binom{M}{2}$ *distancias, y si* M *es un número entero grande la cantidad de distancias por calcular sería bastante grande. Basta con notar que para el Ejemplo 2.2, teniendo el* $(7, 16)$*-código* $\mathcal{C}$ *es necesario calcular* $\binom{16}{2} = \frac{16!}{2!14!} = 120$ *distancias para conseguir la distancia mínima.*

Definición 2.6. *El peso de Hamming $w(x)$ de un vector $x \in \mathbb{F}_2^n$, se define como*

$$w(x) = |\{i : 1 \leqslant i \leqslant n, x_i \neq 0\}|, \tag{2.7}$$

es decir, es el número de posiciones de las coordenadas no cero en x.

Definición 2.7. *El peso mínimo $w(\mathcal{C})$ de un código $\mathcal{C}$, se define como*

$$w(\mathcal{C}) = \text{mín}\{w(x) : x \in \mathcal{C} - \{0\}\}, \tag{2.8}$$

es decir, es el mínimo de los pesos de todas las palabras-código no cero en $\mathcal{C}$.

Ejemplo 2.4. *Consideremos el $(7, 16)$-código $\mathcal{C} \subseteq \mathbb{F}_2^7$ dado en el Ejemplo 2.2, y calculemos el peso de Hamming $w(x)$ para cada palabra-código $x \in \mathcal{C}$, así, como el peso mínimo $w(\mathcal{C})$ del $(7, 16)$-código $\mathcal{C}$, véase la Tabla 9.*

Palabras-código de $\mathcal{C}$	Peso $w(x)$
0 0 0 0 0 0 0	0
0 0 0 1 0 1 1	3
0 0 1 0 1 1 1	4
0 0 1 1 1 0 0	3
0 1 0 0 1 1 0	3
0 1 0 1 1 0 1	4
0 1 1 0 0 0 1	3
0 1 1 1 0 1 0	4
1 0 0 0 1 0 1	3
1 0 0 1 1 1 0	4
1 0 1 0 0 1 0	3
1 0 1 1 0 0 1	4
1 1 0 0 0 1 1	4
1 1 0 1 0 0 0	3
1 1 1 0 1 0 0	4
1 1 1 1 1 1 1	7
Peso mínimo $w(\mathcal{C})$	3

Tabla 9: Peso de Hamming para cada palabra-código $x \in \mathcal{C}$ y el peso mínimo de $\mathcal{C}$.

Observación 2.2. *De las Definiciones 2.4 y 2.6, dados $x, y \in \mathbb{F}_2^n$, se sigue que $d(x, y) = d(x - y, 0) = w(x - y)$, ya que si $d(x, y) = d$, entonces hay d coordenadas en las que x e y difieren y $n - d$ coordenadas en las que x e y coinciden, luego, en la diferencia $x - y$ hay $n - d$ ceros y d coordenadas distintas de cero, así, $w(x - y) = d$.*

Definición 2.8. *Definimos la intersección de vectores binarios $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ como el vector*

$$x \cap y = (x_1 y_1, \ldots, x_n y_n), \tag{2.9}$$

el cual tiene un 1 en la i-ésima posición si y sólo si x e y tienen un 1 en la i-ésima posición.

Lema 2.1. *Si $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{F}_2^n$ entonces*

$$w(x + y) = w(x) + w(y) - 2w(x \cap y). \tag{2.10}$$

Demostración. Sean $w(x) = p$ y $w(y) = q$. Supóngase, sin pérdida de generalidad, que $0 \leqslant p \leqslant q \leqslant n$. Si $p = 0 = q$ entonces $x = 0 = y$ y el resultado es trivialmente cierto. Si $p = 0$ y $p < q$ entonces $x = 0$ y $x + y = y$, de ahí que, $w(x + y) = q$, además $x \cap y = 0$ y $w(x \cap y) = 0$, por consiguiente, $w(x + y) = q = 0 + q - 2 \cdot 0 = w(x) + w(y) - 2w(x \cap y)$, es decir, $w(x + y) = w(x) + w(y) - 2w(x \cap y)$. Si $0 < p \leqslant q$ y $r = w(x \cap y)$ entonces x e y coinciden en r coordenadas, de ahí que, $r \leqslant p \leqslant q$. Hay $p - r$ coordenadas con $1's$ en x, ninguna de las cuales coincide con $q - r$ coordenadas con $1's$ en y. Entonces, al sumar x e y se obtienen r coordenadas con $0's$ y hay $(p - r) + (q - r)$ coordenadas con $1's$, las coordenadas restantes tienen $0's$, así, $w(x + y) = p + q - 2r$, por otro lado, $w(x) + w(y) - 2w(x \cap y) = p + q - 2r$, por lo tanto, $w(x + y) = w(x) + w(y) - 2w(x \cap y)$, con lo cual queda establecido el resultado. □

El siguiente resultado establece que la distancia de Hamming es una métrica.

Teorema 2.1. *La función distancia de Hamming* $d : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{N} \cup \{0\}$ *dada por* $d(x,y)$, *satisface las siguientes propiedades para toda* $x, y, z \in \mathbb{F}_2^n$,

i) $d(x,y) \geqslant 0$, *y* $d(x,y) = 0$ *si y sólo si* $x = y$.

ii) $d(x,y) = d(y,x)$.

iii) $d(x,y) \leqslant d(x,z) + d(z,y)$.

Por lo tanto, el par $(\mathbb{F}_2^n, d)$ *es un espacio métrico.*

Demostración. Sean $x, y, z \in \mathbb{F}_2^n$.

i) Si $x = y$ entonces $d(x,y) = 0$. Si $x \neq y$ entonces x e y difieren en al menos una coordenada, de ahí que, $d(x,y) \geqslant 1 > 0$. De cualquier forma, para cualquier x, y en $\mathbb{F}_2^n$, $d(x,y) \geqslant 0$. Ahora bien, si $d(x,y) = 0$, esto significa que hay cero coordenadas en las que x e y difieren, es decir, $x = y$. Por consiguiente, $d(x,y) = 0$ si y sólo si $x = y$.

ii) Ahora, $d(x,y) = |\{i : 1 \leqslant i \leqslant n, x_i \neq y_i\}| = |\{i : 1 \leqslant i \leqslant n, y_i \neq x_i\}| = d(y,x)$, es decir, $d(x,y) = d(y,x)$.

iii) Finalmente, como $w(x-y) = w(x+y)$ y $2z = 0$ tenemos que:

$$\begin{aligned} d(x,y) &= w(x-y) \\ &= w(x+y) \\ &= w(x+2z+y) \\ &= w((x+z)+(z+y)) \\ &= w(x+z)+w(z+y)-2w((x+z)\cap(z+y)) \\ &\leqslant w(x+z)+w(z+y) \\ &= w(x-z)+w(z-y) \\ &= d(x,z)+d(z,y) \end{aligned}$$

de ahí que, $d(x,y) \leqslant d(x,z) + d(z,y)$.

Por lo tanto, d es una distancia, llamada distancia de Hamming y $(\mathbb{F}_2^n, d)$ resulta un espacio métrico. □

2.1.2 Códigos lineales

Definición 2.9. *Un código* $\mathcal{C} \subseteq \mathbb{F}_2^n$ *es un código lineal binario si es un subespacio del espacio vectorial* $\mathbb{F}_2^n$.

Comentario 2.4. *Indistintamente usaremos el concepto de código lineal binario o código lineal sobre* $\mathbb{F}_2$.

Definición 2.10. *Sea* $H \in M_{(n-k)\times n}(\mathbb{F}_2)$ *arbitraria. Llamamos código lineal binario* $\mathcal{C}$ *con matriz* H *al conjunto que consiste de todos los vectores* $x \in \mathbb{F}_2^n$ *tales que* $Hx^T = 0$, *es decir,*

$$\mathcal{C} = \{x \in \mathbb{F}_2^n : Hx^T = 0\}. \tag{2.11}$$

Observación 2.3. *De la Definición 2.10 veamos que el conjunto* $\mathcal{C}$ (2.11) *es un subespacio de* $\mathbb{F}_2^n$. *Dados* $x, y \in \mathcal{C}$ *y* $\lambda \in \mathbb{F}_2$ *entonces* $H(x+y)^T = Hx^T + Hy^T = 0+0 = 0$ *y* $H(\lambda x)^T = H\lambda x^T = \lambda Hx^T = \lambda 0 = 0$, *es decir,* $H(x+y)^T = 0$ *y* $H(\lambda x)^T = 0$ *para cada* $x, y \in \mathcal{C}$ *y* $\lambda \in \mathbb{F}_2$, *por consiguiente, si* $x, y \in \mathcal{C}$ *y* $\lambda \in \mathbb{F}_2$ *entonces* $x + y \in \mathcal{C}$ *y* $\lambda x \in \mathcal{C}$. *Así, queda demostrado que* $\mathcal{C}$ *es un subespacio de* $\mathbb{F}_2^n$.

Ejemplo 2.5. *El* $[7,4]$*-código de Hamming estudiado en la sección 1.3 del capítulo 1 es un* $(7,16)$*-código lineal binario y es subespacio vectorial de* $\mathbb{F}_2^7$. *Otro ejemplo, es el* $[8,5]$*-código LDPC que es un* $(8,32)$*-código lineal binario y es subespacio vectorial de* $\mathbb{F}_2^8$ *estudiado en el Ejemplo 4.1. Un ejemplo más, es el* $[9,4]$*-código LDPC que es un* $(9,16)$*-código lineal binario y es subespacio vectorial de* $\mathbb{F}_2^9$ *estudiado en el Ejemplo 4.2.*

Observación 2.4. *En la misma Definición 2.10 de código lineal binario, mencionamos una matriz H del código $\mathcal{C}$, es importante señalar que de acuerdo a la Observación 2.3, $\mathcal{C}$ es el espacio vectorial de las soluciones de la ecuación matricial $Hx^T = 0$, esto es, dado $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$ y la matriz $H = (h_{ij}) \in M_{(n-k)\times n}(\mathbb{F}_2)$, entonces $Hx^T = 0$ es equivalente a*

$$\begin{array}{ccccccccc} h_{11}x_1 & + & h_{12}x_2 & + & \cdots & + & h_{1n}x_n & = & 0 \\ h_{21}x_1 & + & h_{22}x_2 & + & \cdots & + & h_{2n}x_n & = & 0 \\ & & & & & & & \vdots & \\ h_{n-k,1}x_1 & + & h_{n-k,2}x_2 & + & \cdots & + & h_{n-k,n}x_n & = & 0 \end{array}$$

por lo tanto, los renglones de H son los coeficientes de un sistema de ecuaciones lineales homogéneas y las soluciones son precisamente las palabras-código en $\mathcal{C}$. Estas ecuaciones lineales son llamadas ecuaciones de chequeo de paridad, ya que un chequeo de paridad en un código lineal binario $\mathcal{C}$ es $hx^T = 0$ para cada palabra-código $x \in \mathcal{C}$ y cada renglón h de la matriz H. De este modo H es conocida como la matriz de chequeo de paridad del código $\mathcal{C}$.

Definición 2.11. *Sea $H \in M_{(n-k)\times n}(\mathbb{F}_2)$. H es llamada la matriz de chequeo de paridad de un código lineal binario $\mathcal{C}$, si H es la matriz asociada al sistema de ecuaciones lineales homogéneas cuyas soluciones determinan a $\mathcal{C}$.*

Otra matriz relacionada con el código lineal binario $\mathcal{C}$, es la matriz generadora, ya que desde que $\mathcal{C}$ es un espacio vectorial, lo podemos describir proporcionando una base.

Definición 2.12. *Sea $\mathcal{C}$ un código lineal sobre $\mathbb{F}_2$ y $G \in M_{k\times n}(\mathbb{F}_2)$. Una matriz G cuyo espacio renglón es igual a $\mathcal{C}$ es llamada una matriz generadora para $\mathcal{C}$. Recíprocamente, si G es una matriz con entradas en $\mathbb{F}_2$, su espacio renglón es llamado el código lineal binario generado por G.*

Observación 2.5. *Si $\mathcal{C}$ es un código lineal binario con matriz generadora G, entonces las palabras-código en $\mathcal{C}$ son precisamente las combinaciones lineales de los renglones de G. Esto es,*

$$\mathcal{C} = \{x \in \mathbb{F}_2^n : x = uG; u \in \mathbb{F}_2^k\} \tag{2.12}$$

Esto proporciona un método muy simple para codificar los dígitos del mensaje fuente u, ya que los renglones de G forman una base para $\mathcal{C}$. En general, diversas matrices generadoras G describen el mismo código lineal binario $\mathcal{C}$, pues recordemos que la base de un espacio vectorial no es única.

Comentario 2.5. *La matriz de chequeo de paridad H de un código lineal binario $\mathcal{C}$ no es única, ya que al realizar operaciones elementales sobre los renglones tendríamos matrices equivalentes por renglones, que seguirían determinando el mismo conjunto de soluciones. Similarmente ocurre con la matriz generadora G de un código lineal binario $\mathcal{C}$.*

Si usamos la codificación sistemática obtenemos una palabra-código x en donde las primeras k componentes de x son iguales a los bits del mensaje fuente u, seguida por $n-k$ bits de redundancia, y al usar esta codificación podemos hallar a la matriz generadora y a la matriz de chequeo de paridad en su forma estándar.

Definición 2.13. *Una matriz generadora G de un código lineal binario $\mathcal{C}$ está en la forma estándar si*

$$G = \left[I_k \quad | \quad A^T\right], \tag{2.13}$$

donde I_k es la matriz identidad de tamaño $k \times k$ y A^T es una matriz de tamaño $k \times (n-k)$ con entradas en $\mathbb{F}_2$. Decimos que G es la matriz generadora estándar del código lineal binario $\mathcal{C}$.

Definición 2.14. *Una matriz de chequeo de paridad H de un código lineal binario $\mathcal{C}$ está en la forma estándar si*

$$H = \left[A \quad | \quad I_{n-k}\right], \tag{2.14}$$

donde A es una matriz de tamaño $(n-k) \times k$ con entradas en $\mathbb{F}_2$ e I_{n-k} es la matriz identidad de tamaño $(n-k) \times (n-k)$. Decimos que H es la matriz de chequeo de paridad estándar del código lineal binario $\mathcal{C}$.

Proposición 2.1. *Si $\mathcal{C}$ es un código lineal binario con matriz de chequeo de paridad $H = \left[A \quad | \quad I_{n-k}\right] \in M_{(n-k)\times n}(\mathbb{F}_2)$, entonces su matriz generadora es $G = \left[I_k \quad | \quad A^T\right]$ y viceversa.*

Demostración. Se puede revisar la demostración en [LA14] pág. 11 – 12. □

Proposición 2.2. *Si $\mathcal{C}$ es un código lineal sobre $\mathbb{F}_2$ con matriz de chequeo de paridad $H = \begin{bmatrix} A & | & I_{n-k} \end{bmatrix} \in M_{(n-k)\times n}(\mathbb{F}_2)$, entonces $\dim \mathcal{C} = k$ y $|\mathcal{C}| = 2^k$.*

Demostración. Dado que $\mathrm{rango}(H) = n - k$, $\dim \mathbb{F}_2^n = n$ y conocemos el resultado que establece que $\dim \mathbb{F}_2^n = \mathrm{nulidad}(H) + \mathrm{rango}(H)$, de esto se sigue que

$$\begin{aligned} \mathrm{nulidad}(H) &= \dim \mathbb{F}_2^n - \mathrm{rango}(H) \\ &= n - (n-k) \\ &= n - n + k \\ &= k. \end{aligned}$$

Ahora, debido a que $\mathcal{C} = \{x \in \mathbb{F}_2^n : Hx^T = 0\}$, tenemos $\dim \mathcal{C} = \mathrm{nulidad}(H) = k$. Además, como toda palabra-código x se puede escribir como combinación lineal de k vectores en $\mathcal{C}$, es decir, $x = \sum_{i=1}^{n} \alpha_i x_i$ donde $\alpha_i \in \mathbb{F}_2$, y $x_i \in \mathcal{C}$ con $i = 1, \ldots, k$. Entonces, hay 2^k posibles combinaciones lineales, luego $|\mathcal{C}| = 2^k$. □

Teorema 2.2. *La distancia mínima de un código lineal binario $\mathcal{C}$ es igual al peso mínimo de cada palabra-código diferente de cero.*

Demostración. Sea $d(\mathcal{C})$ la distancia mínima de un código lineal $\mathcal{C}$, es decir,

$$\begin{aligned} d(\mathcal{C}) &= \min\{d(x,y) : x, y \in \mathcal{C}, x \neq y\} \\ &= \min\{d(x-y, 0) : x, y \in \mathcal{C}, x \neq y\} \\ &= \min\{w(x-y) : x, y \in \mathcal{C}, x \neq y\}, \end{aligned}$$

como $x, y \in \mathcal{C}$ y $x \neq y$ entonces $z := x - y \in \mathcal{C}$ y $z \neq 0$, de ahí que, $\min\{w(x-y) : x, y \in \mathcal{C}, x \neq y\} = \min\{w(z) : z \in \mathcal{C}, z \neq 0\} = w(\mathcal{C})$, por lo tanto, $d(\mathcal{C}) = w(\mathcal{C})$. □

Teorema 2.3. *([Rom92], pág. 95-96) La tasa de un (n, M)-código lineal $\mathcal{C}$ sobre el campo $\mathbb{F}_2$ es*

$$R = \frac{1}{n} \log_2 M. \tag{2.15}$$

Demostración. Sea $\mathcal{C}$ un (n, M)-código lineal binario con $M = |\mathcal{C}| = 2^k$, luego, $M = 2^k$. Entonces

$$\begin{aligned} M = 2^k &\Leftrightarrow \log M = \log 2^k \\ &\Leftrightarrow \log M = k \cdot \log 2 \\ &\Leftrightarrow k = \frac{\log M}{\log 2} = \log_2 M. \end{aligned} \tag{2.16}$$

Ahora, sabemos que la tasa de un código es

$$\begin{aligned} R &= \frac{k}{n} \quad \text{(por Definición 2.3 de tasa)} \\ &= \frac{\log_2 M}{n} \quad \text{(por (2.16)).} \end{aligned}$$

Por lo tanto, la tasa del (n, M)-código lineal binario $\mathcal{C}$ es $R = \frac{1}{n} \log_2 M$. □

Ejemplo 2.6. *La tasa del $(7, 16)$-código binario $\mathcal{C}$ definido en el Ejemplo 2.2 es $R = \frac{1}{7} \log_2 16 = \frac{1}{7} \cdot \frac{\log 16}{\log 2} = \frac{1}{7} \cdot 4 = \frac{4}{7}$. Ahora, este código es el $[7, 4]$-código de Hamming estudiado en la sección 1.3 del capítulo 1, en donde se obtuvo que $k = 4$ y $n = 7$, así, de acuerdo a la Definición 2.3 de tasa de un código vemos que $R = \frac{k}{n} = \frac{4}{7}$. Con esto intentamos mostrar que de acuerdo a ciertos parámetros conocidos es posible determinar la tasa del código.*

Comentario 2.6. *Denotemos por $[n, k, d]$ los parámetros de un código lineal $\mathcal{C}$ donde n es la longitud de las palabras-código o simplemente la longitud del código, k es la dimensión del código lineal $\mathcal{C}$ ($\dim \mathcal{C} = k$), y d es la distancia mínima del código lineal $\mathcal{C}$ ($d(\mathcal{C}) = d$). Entonces, la notación usual para describir a un código lineal $\mathcal{C}$ de longitud n, dimensión k y distancia mínima d será $[\mathbf{n}, \mathbf{k}, \mathbf{d}]$-**código lineal** $\mathcal{C}$, en caso de no conocer la distancia mínima, la notación será $[\mathbf{n}, \mathbf{k}]$-**código lineal** $\mathcal{C}$.*

Para cada código lineal binario $\mathcal{C}$ asociamos el código dual $\mathcal{C}^{\perp}$.

Definición 2.15. *Sea $\mathcal{C}$ un código lineal binario. El conjunto*

$$\mathcal{C}^{\perp} = \{v \in \mathbb{F}_2^n : xv^T = 0, \forall x \in \mathcal{C}\}, \tag{2.17}$$

es llamado el código dual de $\mathcal{C}$.

Comentario 2.7. *El código dual $\mathcal{C}^{\perp}$ de un código lineal $\mathcal{C}$, también es un código lineal.*

Teorema 2.4. *i) Si G es una matriz generadora para $\mathcal{C}$, entonces*

$$\mathcal{C}^{\perp} = \{v \in \mathbb{F}_2^n : Gv^T = 0\}. \tag{2.18}$$

ii) El dual $\mathcal{C}^{\perp}$ de un código lineal es un $[n, n-k]$-código lineal.

iii) Para cualquier código lineal $\mathcal{C}$, tenemos $\mathcal{C}^{\perp\perp} = \mathcal{C}$.

Demostración. Se puede revisar la demostración en [Rom92] pág. $199 - 200$. □

Observación 2.6. *De acuerdo al Teorema 2.4, podemos decir que el código dual $\mathcal{C}^{\perp}$ está caracterizado por*

$$\mathcal{C}^{\perp} = \{v \in \mathbb{F}_2^n : v = uH, u \in \mathbb{F}_2^{n-k}\} = \{v \in \mathbb{F}_2^n : Gv^T = 0\}.$$

En la misma forma que el código lineal $\mathcal{C}$ por

$$\mathcal{C} = \{x \in \mathbb{F}_2^n : x = uG, u \in \mathbb{F}_2^k\} = \{x \in \mathbb{F}_2^n : Hx^T = 0\}.$$

Esto es, si H es una matriz de chequeo de paridad del código lineal $\mathcal{C}$, entonces H es una matriz generadora para el código dual $\mathcal{C}^{\perp}$. Y si G es una matriz generadora del código lineal $\mathcal{C}$, entonces G es una matriz de chequeo de paridad para el código dual $\mathcal{C}^{\perp}$.

2.1.3 Síndrome de decodificación

Un proceso de decodificación eficiente para códigos lineales puede obtenerse a través del uso de la matriz de chequeo de paridad H del código lineal $\mathcal{C}$. Si x es transmitida, x es una palabra-código, luego, $Hx^T = 0$. Si el canal provoca algunos errores, esto es, definamos el vector $e = y - x$ llamado el patrón de error tal que, si $e \neq 0$, entonces es muy probable que $Hy^T \neq 0$ donde $y = x + e$ conocida como la palabra recibida. De esta manera obtenemos el vector $s = Hy^T$ llamado el síndrome, el cual definimos a continuación.

Definición 2.16. *Sea $\mathcal{C}$ un $[n, k]$-código lineal, con matriz de chequeo de paridad $H \in M_{(n-k)\times n}(\mathbb{F}_2)$. Para cualquier $y \in \mathbb{F}_2^n$, la palabra $s = Hy^T$ es llamada el síndrome de y.*

Comentario 2.8. *Es importante mencionar que $y \in \mathcal{C}$ si y sólo si el síndrome de y es 0.*

Observación 2.7. *El síndrome depende sólo del patrón de error e y no de la palabra-código x transmitida, pues si $s = Hy^T = H(x+e)^T = Hx^T + He^T = 0 + He^T = He^T$, es decir, $s = He^T$. El síndrome proporciona cierta información sobre e, pero no la suficiente. Esto es debido a que para un $s \in \mathbb{F}_2^n$ fijo, el conjunto de soluciones de $He^T = s$ forma una clase del código $\mathcal{C}$.*

Definición 2.17. *Sea $(\mathbb{F}_2^n, +)$ un grupo, $\mathcal{C} \leqslant \mathbb{F}_2^n$, para cada $a, b \in \mathbb{F}_2^n$ decimos que a es congruente a b módulo $\mathcal{C}$, lo cual denotamos $a \equiv b \pmod{\mathcal{C}}$ si y sólo si $a - b \in \mathcal{C}$.*

Lema 2.2. *La relación $a \equiv b \pmod{\mathcal{C}}$ es una relación de equivalencia.*

Demostración. Veamos que para cada $a, b, c \in \mathbb{F}_2^n$ se cumple lo siguiente:

i) $a \equiv a \pmod{\mathcal{C}}$.
Como $a - a = 0 \in \mathcal{C}$ tenemos que $a \equiv a \pmod{\mathcal{C}}$.

ii) $a \equiv b \pmod{\mathcal{C}}$ implica que $b \equiv a \pmod{\mathcal{C}}$.
Supóngase que $a \equiv b \pmod{\mathcal{C}}$ entonces $a - b \in \mathcal{C}$, luego, como $\mathcal{C}$ es un grupo, $-(a - b) \in \mathcal{C}$, eso implica que $b - a \in \mathcal{C}$, de ahí que, $b \equiv a \pmod{\mathcal{C}}$.

iii) $a \equiv b \pmod{\mathcal{C}}$ y $b \equiv c \pmod{\mathcal{C}}$ implica que $a \equiv c \pmod{\mathcal{C}}$.
Supóngase que $a \equiv b \pmod{\mathcal{C}}$ y $b \equiv c \pmod{\mathcal{C}}$ entonces $a - b \in \mathcal{C}$ y $b - c \in \mathcal{C}$, luego, $(a - b) + (b - c) \in \mathcal{C}$, es decir, $a - c \in \mathcal{C}$, por consiguiente, $a \equiv c \pmod{\mathcal{C}}$.

Por lo tanto, $a \equiv b \pmod{\mathcal{C}}$ es una relación de equivalencia en $\mathbb{F}_2^n$. □

Definición 2.18. *Sea la relación de equivalencia* $a \equiv b \pmod{\mathcal{C}}$ *en* $\mathbb{F}_2^n$*, definimos la clase de* a*, lo cual se denota por* $\bar{a}$ *o* $[a]$*, como*

$$[a] = \{x \in \mathbb{F}_2^n : a \equiv x \pmod{\mathcal{C}}\}. \tag{2.19}$$

Proposición 2.3. *Para cada* $a \in \mathbb{F}_2^n$, $[a] = a + \mathcal{C}$.

Demostración. Sea $a \in \mathbb{F}_2^n$.

$$\begin{aligned}
[a] &= \{b \in \mathbb{F}_2^n : b \equiv a \pmod{\mathcal{C}}\} \\
&= \{b \in \mathbb{F}_2^n : b - a \in \mathcal{C}\} \\
&= \{b \in \mathbb{F}_2^n : b - a = x \text{ para algún } x \in \mathcal{C}\} \\
&= \{b \in \mathbb{F}_2^n : b = a + x \text{ para algún } x \in \mathcal{C}\} \\
&= a + \mathcal{C}.
\end{aligned}$$

Por lo tanto, $[a] = a + \mathcal{C}$. □

Proposición 2.4. *Sea* $\mathcal{C}$ *un* $[n, k]$*-código lineal binario. Entonces*

i) Todo vector $b \in \mathbb{F}_2^n$ *está en alguna clase.*

ii) Dos vectores a *y* b *están en la misma clase si y sólo si* $a - b \in \mathcal{C}$.

iii) Cada clase contiene 2^k *vectores.*

Demostración. Sea $\mathcal{C}$ un $[n, k]$-código lineal binario.

i) Sea $b \in \mathbb{F}_2^n$, $b = b + 0 \in b + \mathcal{C}$, ya que $0 \in \mathcal{C}$, es decir, $b \in b + \mathcal{C}$. Así, todo vector $b \in \mathbb{F}_2^n$ está en alguna clase.

ii) Supóngase que $a, b \in u + \mathcal{C}$ para algún $u \in \mathbb{F}_2^n$ entonces $a = u + x_1$ y $b = u + x_2$ para algunos $x_1, x_2 \in \mathcal{C}$ luego $a - b = (u + x_1) - (u + x_2) = x_1 - x_2 \in \mathcal{C}$, es decir, $a - b \in \mathcal{C}$. Recíprocamente, si $a - b \in \mathcal{C}$ entonces $a - b = x$ para algún $x \in \mathcal{C}$, luego $a = b + x \in b + \mathcal{C}$, de ahí que, $a \in b + \mathcal{C}$, por consiguiente, $a, b \in b + \mathcal{C}$.

iii) Finalmente, como $\mathcal{C}$ tiene 2^k palabras-código distintas entonces $a + \mathcal{C}$ tiene 2^k vectores distintos, ya que si $a + x_1 = a + x_2$ para $x_1 \neq x_2$ en $\mathcal{C}$ entonces $x_1 = x_2$, lo cual es absurdo.

□

Comentario 2.9. *La relación de equivalencia* $a \equiv b \pmod{\mathcal{C}}$ *en* $\mathbb{F}_2^n$ *induce una partición de* $\mathbb{F}_2^n$ *en clases de equivalencia no vacías y disjuntas por parejas, teniendo así, el conjunto de todas las clases de equivalencia* $\mathbb{F}_2^n / \equiv = \{[a] : a \in \mathbb{F}_2^n\}$.

Observación 2.8. *Como* $|\mathbb{F}_2^n| = 2^n$, $|a + \mathcal{C}| = 2^k$ *y* $\mathbb{F}_2^n = \cup\{a + \mathcal{C} : a \in \mathbb{F}_2^n\}$*, tenemos que,* $2^n = |\cup\{a + \mathcal{C} : a \in \mathbb{F}_2^n\}| = r2^k$ *donde* r *es el número de clases de equivalencia en* $\mathbb{F}_2^n$*, luego* $2^n = r2^k$*, de ahí que,* $r = 2^{n-k}$*. Hay* 2^{n-k} *clases de* $\mathcal{C}$*, correspondientes a los* 2^{n-k} *posibles síndromes* s*. Así, una vez que el receptor calcula* s*, reduce su búsqueda para* e *de* 2^n *a* 2^k *posibilidades, a saber, los elementos de la clase correspondiente a* s.

Comentario 2.10. *Al considerar* $V = \mathbb{F}_2^n$ *y* $W = \mathcal{C}$*, con* $\mathcal{C} \leqslant V$*, construimos* V/W *el conjunto cociente que al darle la estructura de espacio vectorial es un espacio cociente con la operación suma definida como* $(a + W) + (b + W) = (a + b) + W$ *y la operación multiplicación por un escalar dada por* $\lambda(a + W) = \lambda a + W$.

En el siguiente teorema daremos algunas propiedades del síndrome de decodificación.

Teorema 2.5. *Para un* $[n, k]$*-código lineal binario* $\mathcal{C}$ *con matriz de chequeo de paridad* $H \in M_{(n-k)\times n}(\mathbb{F}_2)$*, sea* $s = Hy^T$ *el síndrome de la palabra recibida* y*. Entonces*

i) s *es un vector columna de longitud* $n-k$.

ii) *Si* $e=(e_1,e_2,\ldots,e_n)\in\mathbb{F}_2^n$ *y* $e_{i_j}=1$ *para* $i_j\in\{1,2,\ldots,n\}$, $j=1,\ldots,t$, *es tal que* $y=x+e$ *entonces*

$$s=\sum_{j=1}^{t}H_{i_j} \tag{2.20}$$

donde H_{i_j} *es la* i_j*-ésima columna de* H, *es decir, el síndrome* s *es igual a la suma de las columnas de* H *en donde los errores ocurren.*

iii) *Dos vectores están en la misma clase de* $\mathcal{C}$ *si y sólo si ellos tienen el mismo síndrome.*

iv) *Hay una correspondencia uno a uno entre los síndromes y las clases.*

Demostración. i) La primera propiedad es inmediata a partir de la Definición 2.16 de síndrome.

ii) Sea $\{v_i\}_{i=1}^n$ la base canónica de $\mathbb{F}_2^n$, dado $e\in\mathbb{F}_2^n$ tenemos que $e=\sum_{i=1}^n e_iv_i$ donde $e_i\in\mathbb{F}_2$. Entonces

$$\begin{aligned}
s &= Hy^T\\
&= He^T \quad \text{(por Observación 2.7)}\\
&= H(\sum_{i=1}^{n}e_iv_i)^T\\
&= H(\sum_{j=1}^{t}e_{i_j}v_{i_j})^T\\
&= \sum_{j=1}^{t}e_{i_j}Hv_{i_j}^T\\
&= \sum_{j=1}^{t}H_{i_j}
\end{aligned}$$

donde H_{i_j} es la i_j-ésima columna de H.

iii) Sean u_1 y u_2 dos vectores en $\mathbb{F}_2^n$. Denotamos como s_1 y s_2 los síndromes de u_1 y u_2 respectivamente.

$$\begin{aligned}
u_1,u_2\in u_1+\mathcal{C} &\Leftrightarrow u_1-u_2\in\mathcal{C}\\
&\Leftrightarrow H(u_1-u_2)^T=0\\
&\Leftrightarrow Hu_1^T=Hu_2^T\\
&\Leftrightarrow s_1=s_2.
\end{aligned}$$

iv) La última propiedad se sigue de que como hay 2^{n-k} clases distintas hay 2^{n-k} síndromes distintos.

□

2.1.4 Código detector-corrector de errores

Al diseñar un algoritmo de decodificación lo ideal sería que se lograra detectar y corregir todos los errores producidos en el canal; sin embargo, la cantidad de errores detectados y corregidos no siempre serán iguales, ya que en ocasiones no es posible corregir todos los errores detectados. En sí, necesitamos una probabilidad de decodificación correcta muy cercana a uno, y además, tampoco se garantiza detectar todos los errores, pues la probabilidad de no detectar el error debería ser muy cercana a cero.

Sea $F\subseteq\mathbb{F}_2^n$, considérese a F como el conjunto de patrones de error e que tienen probabilidad moderada de ocurrir en el canal. Ahora, sea E el subconjunto de patrones de error e en F que tienen probabilidad alta de ocurrir. Dado un código lineal $\mathcal{C}$, deseamos diseñar, si es posible, un decodificador que detectará los patrones de error e en F y corregirá los patrones de error e en E.

Definición 2.19. *El código $\mathcal{C}$ es llamado E-corrector, F-detector, si es posible diseñar un decodificador para $\mathcal{C}$ tal que para cada patrón de error e, si $e \in E$, e será corregido, y si $e \in F$, e será detectado o corregido.*

La distancia mínima del código juega un papel esencial en la respuesta a la pregunta ¿cuántos errores puede corregir un código?, pero antes de responder dicha pregunta damos la siguiente definición.

Definición 2.20. *La esfera (o esfera de Hamming) de radio r y centro u, lo denotamos por $B_r(u)$, se define como*

$$B_r(u) = \{v \in \mathbb{F}_2^n : d(u,v) \leqslant r\}. \tag{2.21}$$

Teorema 2.6. *Un código $\mathcal{C}$ con distancia mínima d (o peso mínimo d) puede corregir $[\frac{1}{2}(d-1)]$ ó menos errores.*

Demostración. Sea $t = [\frac{1}{2}(d-1)]$ (t es el mayor entero menor o igual a $\frac{1}{2}(d-1)$), si una palabra-código x es transmitida y ocurren t ó menos errores, la palabra recibida y se encontrará en la esfera de radio t alrededor de la palabra-código transmitida x. Para que el código pueda corregir t errores o menos verifiquemos que las esferas de radio t con centro en las palabras-código son disjuntas. Sean $x_1, x_2 \in \mathcal{C}$ y supóngase que $B_t(x_1) \cap B_t(x_2) \neq \emptyset$, entonces existe $v \in \mathbb{F}_2^n$ tal que $v \in B_t(x_1)$ y $v \in B_t(x_2)$, por consiguiente, $d(x_1,x_2) \leqslant d(x_1,v) + d(v,x_2) \leqslant t + t = 2t$, es decir, $d(x_1,x_2) \leqslant 2t$ pero $d \leqslant d(x_1,x_2) \leqslant 2t$ entonces $d \leqslant 2t$. Dado que $t \leqslant \frac{1}{2}(d-1)$, puesto que $t = [\frac{1}{2}(d-1)]$, tenemos que $2t \leqslant d-1$ luego $d \leqslant 2t \leqslant d-1$, de ahí que, $d \leqslant d-1$, lo cual es absurdo. En consecuencia, las esferas de radio t con centro en las palabras-código son disjuntas. De manera que, la palabra recibida y está más cerca de x que de cualquier otra palabra-código u. Así, la decodificación de vecino más cercano corregirá estos errores. □

2.2 TEOREMA DE CODIFICACIÓN DE CANAL CON RUIDO DE SHANNON

Una vez discutidos algunos conceptos escenciales sobre los códigos lineales, en el sentido de la codificación y decodificación, platicaremos un poco sobre lo que ocurre en el proceso intermedio de la comunicación, es decir, la transmisión de la palabra-código sobre el canal. Algo que nos interesa saber es, ¿cuál es la tasa máxima alcanzada sobre un canal dado para la transmisión de un mensaje? La respuesta fue probada por Claude Shannon en 1948, en el Teorema de Codificación de Canal con Ruido, consultar [Sha48].

La intención de esta sección es únicamente enunciar el Teorema de Shannon sin demostraciones, ya que este teorema asegura la existencia de un código a partir de algunas condiciones y el cual tendrá ciertos parámetros. Para enunciar el teorema, debemos introducir primero el concepto de entropía, una medida de incertidumbre para los valores asumidos por las variables aleatorias, pero antes la siguiente definición.

Definición 2.21. *Un canal discreto sin memoria consiste de símbolos de entrada de un alfabeto X, símbolos de salida de un alfabeto Y, y un conjunto de probabilidades del canal $p(y_j|x_i)$, que satisface*

$$\sum_{j=1}^{t} p(y_j|x_i) = 1 \tag{2.22}$$

para toda i y donde $p(y_j|x_i)$ es la probabilidad condicional de y_j dado x_i.

Definición 2.22. *La entropía de una variable aleatoria discreta X, $H(X)$, es*

$$H(X) = \sum_{x \in X} p(x) \log \frac{1}{p(x)}. \tag{2.23}$$

Comentario 2.11. *A su vez, son definidas las entropía conjunta $H(X,Y)$ de dos variables aleatorias discretas X y Y, y la entropía condicional $H(Y|X)$, como:*

$$H(X,Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{1}{p(x,y)},$$

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X = x) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{1}{p(y|x)},$$

respectivamente. Además $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.

Definición 2.23. *La información mutua entre dos variables aleatorias* X *y* Y, $I(X; Y)$, *es*

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x) p(y)}. \tag{2.24}$$

Observación 2.9. *Notar que* $I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$. *Al suponer un canal binario, el máximo valor de* $H(X)$ *es* 1, *con las unidades expresadas en bits. Luego, si hay ruido la correspondiente reducción en la información mutua refleja la pérdida de la información durante la transmisión.*

Ahora, la capacidad C de un canal es la máxima cantidad de información que puede ser enviada a través del canal, el máximo es tomado sobre todas las distribuciones de entrada.

Definición 2.24. *La capacidad de un canal es la máxima información mutua* $I(X; Y)$, *tomada sobre todas las distribuciones de entrada* $p(x_i)$ *de* X, *esto es,*

$$C = \max_{p(x_i)} I(X; Y). \tag{2.25}$$

El problema de determinar la capacidad de un canal arbitrario es díficil. Sin embargo, es posible determinar las capacidades de algunos tipos especiales de canales, incluyendo el canal simétrico, sobre el cual nuestro trabajo está encaminado.

Teorema 2.7. *La capacidad de un canal simétrico es*

$$C = \log t - \sum_{j=1}^{t} p(y_j|x_i) \log \frac{1}{p(y_j|x_i)} \tag{2.26}$$

para cada $i = 1, \ldots, s$. *Además, la capacidad es lograda por la distribución de entrada uniforme* $p(x_i) = \frac{1}{s}$.

Demostración. Podemos revisar la demostración en [Rom92] pág. 85. □

Corolario 2.1. *La capacidad del canal simétrico binario con probabilidad* f *de recibir incorrectamente el bit y con probabilidad* $1 - f$ *de recibirlo correctamente, es*

$$C = 1 - H(f) \tag{2.27}$$

donde $H(f)$ *es la función entropía.*

La idea del Teorema de Shannon es que, asociado con cada canal discreto sin memoria, hay un número no negativo C llamado la capacidad del canal con la siguiente propiedad. Para cualquier $\epsilon > 0$ y $R < C$, y un n suficientemente grande, existe un código de longitud n y tasa $\geqslant R$, es decir, con al menos 2^{Rn} palabras-código distintas, y un algoritmo de decodificación apropiado, tal que, cuando el código es usado sobre el canal dado, la probabilidad de error de la decodificación es $< \epsilon$.

Teorema 2.8. *Consideremos un canal discreto sin memoria con capacidad* C. *Para algún número positivo* $R < C$, *existe una secuencia* C_n *de códigos* q*-arios, y sus correspondientes algoritmos de decodificación, con las siguientes propiedades:*

i) C_n *es un* $(n, [q^{nR}])$*-código, es decir,* C_n *tiene longitud* n *y tasa de al menos* R. *(*$[q]$ *es la parte entera de* q*.)*

ii) La probabilidad máxima de error del algoritmo de decodificación se aproxima a 0 *cuando* $n \to \infty$, *esto es,*

$$P_e^{max}(n) \to 0. \tag{2.28}$$

Demostración. Podemos consultar la demostración en [Rom92] pág. $107 - 111$. □

Para más detalles sobre los conceptos definidos y el teorema el lector puede consultar las siguientes referencias [Sha48], [Rom92], [McE04] y [WK03].

2.3 ELEMENTOS DE LA TEORÍA DE GRAFOS

La interpretación teórica que se da de los grafos a los códigos detectores-correctores de errores ha conducido a técnicas para la construcción de códigos cuyo desempeño es muy cercano al límite de Shannon establecido en el Teorema 2.8, por lo que nos interesa conocer la estructura básica de los grafos.

Comenzaremos con la definición para un grafo y procederemos a discutir la clase importante de los grafos bipartitos. Además, presentaremos dos métodos para transformar un grafo no bipartito en un grafo bipartito a través del uso de grafos de incidencia arista-vértice y la cubierta doble que más adelante serán de utilidad.

Los textos revisados para el desarrollo de tal sección fueron [BM08], [Die05], [Gri04] y [WK03].

2.3.1 Conceptos básicos de grafos

Definición 2.25. *Sea V un conjunto finito no vacío y sea $E \subseteq [V]^2 = \{X \subseteq V \mid X$ es de cardinalidad 2 $(|X| = 2)\}$. Una gráfica o grafo $\mathcal{G}$ es un par $\mathcal{G} = (V, E)$ consistente del conjunto $V = \{v_1, v_2, \ldots, v_n\}$ de vértices o nodos y del conjunto $E = \{e_1, e_2, \ldots, e_m\}$ de aristas.*

Observación 2.10. *Para evitar ambigüedades asumimos que $E \cap V = \emptyset$.*

Observación 2.11. *A un grafo $\mathcal{G} = (V, E)$ se le conoce como **grafo no dirigido** cuando no nos interesa la dirección de sus aristas. La estructura del grafo $\mathcal{G}$ es que sus aristas son pares no ordenados de vértices, esto es, $e = \{a, b\}$ con $a, b \in V$.*

Observación 2.12. *A un grafo $\mathcal{G} = (V, E)$ se le conoce como **grafo dirigido** o **digrafo** cuando si nos interesa la dirección de sus aristas. La estructura del grafo $\mathcal{G}$ es diferente a la establecida en la Definición 2.25, pues $E \subseteq V \times V$ y sus aristas son parejas ordenadas de vértices, es decir, $e = (a, b)$ con $a, b \in V$ donde a es el vértice inicial de la arista y b es el vértice final de la arista, [Gri04] pág. 514. Nosotros no trabajaremos sobre digrafos, pero se hace la mención debido a que más adelante en uno de los resultados habrá un concepto relacionado a ellos.*

Comentario 2.12. *La forma usual para ilustrar un grafo es, dibujar un punto para cada vértice y unir por medio de una línea dos vértices si estos forman una arista. Entonces los puntos representan a los vértices y las líneas a las aristas, y así obtendremos el dibujo de un grafo.*

Definición 2.26. *Sea $\mathcal{G} = (V, E)$ un grafo. Una arista de $\mathcal{G}$ cuyos vértices extremos son idénticos es llamado lazo, es decir, $e = \{v, v\} \in E$ con $v \in V$. Una arista de $\mathcal{G}$ cuyos vértices extremos son distintos es llamado un enlace. Dos o más enlaces con el mismo par de vértices como extremos son conocidos como aristas múltiples.*

Ejemplo 2.7. *En la Figura 25 mostramos: un lazo donde $V = \{1\}$ y $E = \{\{1, 1\}\}$; un enlace donde $V = \{2, 3\}$ y $E = \{\{2, 3\}\}$; y las aristas múltiples donde $V = \{4, 5\}$ y $E = \{\{4, 5\}, \{4, 5\}, \{4, 5\}\}$.*

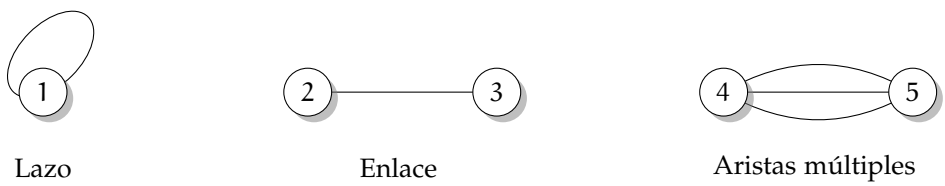


Figura 25: Lazo, enlace y aristas múltiples.

Definición 2.27. *A un grafo $\mathcal{G} = (V, E)$ se le conoce como **grafo simple** si éste no tiene lazos o aristas múltiples.*

Comentario 2.13. *La mayor parte de la teoría que relaciona a un grafo dentro de este trabajo está enfocada al estudio de grafos simples y no dirigidos.*

Ejemplo 2.8. *En la Figura 26 ilustramos el grafo $\mathcal{G} = (V, E)$. El grafo $\mathcal{G}$ consiste del conjunto de vértices $V = \{1,2,3,4,5,6\}$ y del conjuto de aristas $E = \{\{1,4\},\{2,4\},\{3,5\},\{4,5\}\}$, tal grafo $\mathcal{G}$ es un grafo no dirigido pues las aristas no contienen dirección, además, de ser un grafo simple pues no tiene aristas múltiples ni lazos.*

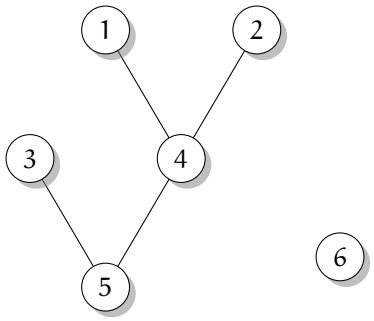


Figura 26: Grafo $\mathcal{G} = (V, E)$.

Descripción de terminología.

- Un vértice es incidente con una arista, si un vértice v forma parte de una arista e, es decir, $v \in e$. Y una arista es incidente con el vértice, si e es una arista del vértice v.
- Dos vértices v_1, v_2 de $\mathcal{G}$ son adyacentes o vecinos, si $e = \{v_1, v_2\}$ es una arista de $\mathcal{G}$. En otras palabras, cuando dos vértices están conectados por una arista se dicen que son adyacentes.

Definición 2.28. *Un multigrafo $\mathcal{G}$ es un par $\mathcal{G} = (V, E)$ de conjuntos disjuntos de vértices y aristas junto con una función que está definida de $E \to V \cup [V]^2$ asignando para cada arista cualesquiera uno o dos vértices como extremos. Así, los multigrafos también pueden tener lazos y aristas múltiples.*

Comentario 2.14. *La Definición 2.28 de multigrafo se enuncia pues en la sección 2.4 se construirán algunos ejemplos que tienen que ver con dicho concepto.*

Definición 2.29. *Sea $\mathcal{G} = (V, E)$ un grafo. Un conjunto de vértices o de aristas en $\mathcal{G}$ es independiente si dos de sus elementos no son adyacentes.*

Definición 2.30. *Sea $\mathcal{G} = (V, E)$ un grafo. Si todos los vértices de $\mathcal{G}$ son parejas adyacentes, entonces $\mathcal{G}$ es un grafo completo.*

Comentario 2.15. *Un grafo completo con n vértices se denota por K^n.*

Ejemplo 2.9. *Sea $\mathcal{G} = (V, E)$ con $V = \{1,2,3,4\}$ y $E = \{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}$, tal grafo es un grafo completo con $n = 4$ vértices, así, $\mathcal{G}$ es un K^4, el cual, ilustramos en la Figura 27.*

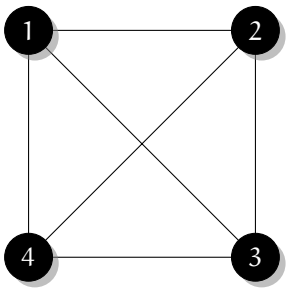


Figura 27: Grafo completo K^4.

Definición 2.31. *Sea $\mathcal{G} = (V, E)$ un grafo. El número de vértices n del grafo $\mathcal{G}$ es su orden, y se denota como $|\mathcal{G}|$ o $|V|$. Y su número de aristas m es denotado por $\|\mathcal{G}\|$ o $|E|$.*

Definición 2.32. *El grado de un vértice v es el número de aristas $|E(v)|$ incidentes sobre un vértice v, esto es igual al número de vecinos de v. El grado de un vértice es denotado por $d_{\mathcal{G}}(v)$ o simplemente $d(v)$.*

Comentario 2.16. *El número de aristas puede ser calculado al sumar todos los grados de los vértices en el grafo $\mathcal{G}$, con lo cual contamos cada arista dos veces, una vez por cada uno de sus vértices, por lo tanto,* $|E| = \frac{1}{2}\sum_{v\in V} d(v)$, *veáse [Die05] pág.* 5.

Definición 2.33. *El número* $\delta(\mathcal{G}) := \min\{d(v) \mid v \in V\}$ *es el grado mínimo de* $\mathcal{G}$.

Definición 2.34. *El número* $\Delta(\mathcal{G}) := \max\{d(v) \mid v \in V\}$ *es el grado máximo de* $\mathcal{G}$.

Definición 2.35. *El número* $d(\mathcal{G}) := \frac{1}{|\mathcal{G}|}\sum_{v\in V} d(v)$ *es el grado promedio del grafo* $\mathcal{G}$.

Observación 2.13. *Claramente se da la siguiente relación:* $\delta(\mathcal{G}) \leqslant d(\mathcal{G}) \leqslant \Delta(\mathcal{G})$.

Comentario 2.17. *El grado promedio cuantifica globalmente lo que está medido localmente por los grados de los vértices, es decir, el número de aristas de* $\mathcal{G}$ *por vértice.*

Ejemplo 2.10. *Para el siguiente grafo* $\mathcal{G} = (V, E)$ *donde* $V = \{1,2,3,4\}$ *y* $E = \{\{1,2\},\{2,3\}\}$, *ilustrado en la Figura 28, calculemos los conceptos que acabamos de definir.*

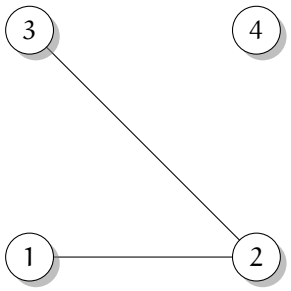


Figura 28: Grafo $\mathcal{G} = (V, E)$ con $V = \{1,2,3,4\}$ y $E = \{\{1,2\},\{2,3\}\}$.

- *El grafo* $\mathcal{G}$ *tiene* $n = 4$ *vértices, así, su orden es* $|\mathcal{G}| = 4$.
- $\mathcal{G}$ *tiene* $m = 2$ *aristas, es decir,* $\|\mathcal{G}\| = 2$.
- *El grado para cada uno de los vértices es:*
  - *Para* $v_1 = 1$ *es* $d(v_1) = 1$.
  - *Para* $v_2 = 2$ *es* $d(v_2) = 2$.
  - *Para* $v_3 = 3$ *es* $d(v_3) = 1$.
  - *Para* $v_4 = 4$ *es* $d(v_4) = 0$, *este vértice es aislado.*
- *El grado mínimo de* $\mathcal{G}$ *es,* $\delta(\mathcal{G}) = 0$.
- *El grado máximo de* $\mathcal{G}$ *es,* $\Delta(\mathcal{G}) = 2$.
- *El grado promedio de* $\mathcal{G}$ *es,* $d(\mathcal{G}) = \frac{1}{|\mathcal{G}|}\sum_{v\in V} d(v) = \frac{1}{4}(d(v_1) + d(v_2) + d(v_3) + d(v_4)) = \frac{1}{4}(1+2+1+0) = \frac{1}{4}\cdot 4 = 1$.

Definición 2.36. *Sea* $\mathcal{G} = (V, E)$ *un grafo. El grafo* $\mathcal{G}$ *es regular si todos los vértices del grafo tienen el mismo grado. Si cada vértice del grafo regular es de grado* r, *diremos que el grafo* $\mathcal{G}$ *es* r*-regular. Y un grafo* $\mathcal{G}$ *es irregular, si éste no es regular.*

Ejemplo 2.11. *Sea el grafo* $\mathcal{G} = (V, E)$ *con* $V = \{1,2,3,4,5,6\}$ *y* $E = \{\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{2,3\},\{2,5\},\{2,6\},\{3,4\},\{3,6\},\{4,5\},\{4,6\},\{5,6\}\}$, *ilustrado en la Figura 29. Si observamos todos los vértices del grafo* $\mathcal{G}$ *vemos que tienen el mismo número de aristas, es decir,* $d(v_i) = 4$ *para todo* $i = 1,2,\ldots,6$, *entonces el grafo* $\mathcal{G}$ *es regular de grado* $r = 4$, *esto es un grafo* $\mathcal{G}$ 4*-regular.*

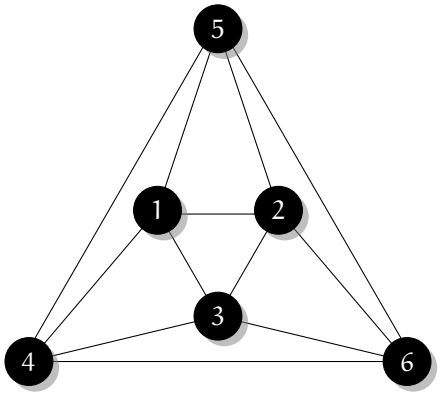


Figura 29: Grafo $\mathcal{G}$ 4-regular.

Definición 2.37. *Un camino simple o trayectoria es un grafo no vacío* $P = (V, E)$ *de la forma* $V = \{v_0, v_1, ..., v_k\}$, $E = \{\{v_0, v_1\}, \{v_1, v_2\}, ..., \{v_{k-1}, v_k\}\}$, *donde los* v_i *son todos distintos. Los vértices* v_0 *y* v_k *están unidos por la trayectoria* P *y son llamados vértices extremos; los vértices* $v_1, ..., v_{k-1}$ *son los vértices interiores de* P.

Observación 2.14. *El número de aristas de un camino simple es su longitud, y el camino simple de longitud* k *es denotado por* P^k. *Frecuentemente nos referimos a trayectoria por la secuencia natural de sus vértices, escribiendo* $P = v_0v_1...v_k$ *y llamamos* P *una trayectoria desde* v_0 *a* v_k.

Ejemplo 2.12. *Ilustramos en la Figura 30 el camino simple* $P = (V, E)$ *con* $V = \{1, 2, 3, 4\}$ *y* $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$. *Como* P *va desde* 1 *a* 4, $P = 1234$, *y debido a que tiene longitud* $k = 3$ *es un camino simple* P^3.

Figura 30: El camino simple o trayectoria P^3.

Definición 2.38. *Si* $P = v_0v_1...v_{k-1}$ *es una trayectoria y* $k \geqslant 3$, *entonces el grafo* $C := P + \{v_{k-1}, v_0\}$ *es llamado un ciclo.*

Observación 2.15. *Usualmente nos referimos a un ciclo por su secuencia cíclica de vértices, es decir,* $C = v_0...v_{k-1}v_0$. *La longitud de un ciclo es su número de aristas (o vértices). Y además, el ciclo de longitud* k *es un* k*-ciclo y es denotado por* C^k.

Ejemplo 2.13. *Ilustramos en la Figura 31 el ciclo* $C = P + \{4, 1\}$ *donde* $P = 1234$, *este ciclo tiene longitud* $k = 4$, *luego, podemos decir que es un ciclo* C^4.

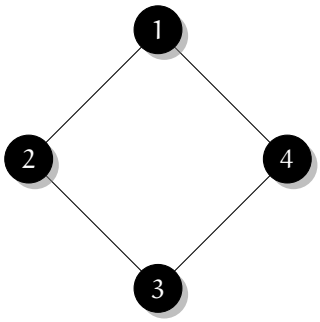


Figura 31: Ciclo C^4.

Ahora hablemos un poco del concepto de conectividad para grafos no dirigidos y digrafos.

Definición 2.39. *Sea $\mathcal{G} = (V, E)$ un grafo no vacío. Decimos que $\mathcal{G}$ es conexo si cualesquiera dos de sus vértices están unidos por una trayectoria en $\mathcal{G}$. Y un grafo $\mathcal{G}$ que no es conexo será llamado disconexo.*

Ejemplo 2.14. *Sea el grafo $\mathcal{G} = (V, E)$ con $V = \{1, 2, 3, 4, 5, 6, 7\}$ y $E = \{\{1, 3\}, \{1, 5\}, \{2, 7\}, \{3, 6\}, \{4, 7\}, \{5, 6\}\}$, ilustrado en la Figura 32. $\mathcal{G}$ es un grafo disconexo, pues no existe una trayectoria entre el vértice 7 y el vértice 3.*

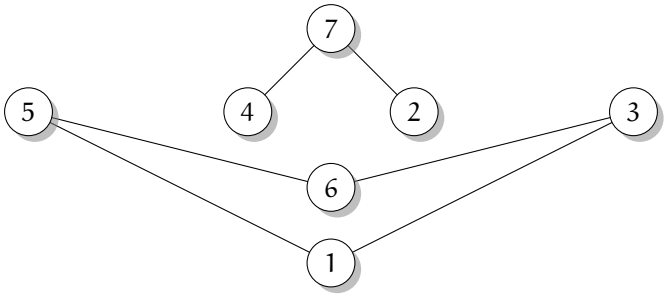


Figura 32: Grafo $\mathcal{G}$ disconexo.

Definición 2.40. *Sea $\mathcal{G} = (V, E)$ un digrafo. Decimos que $\mathcal{G}$ es fuertemente conexo si para cada par de vértices $a, b \in V$ existe una trayectoria dirigida de a hacia b y una trayectoria dirigida de b hacia a.*

Definición 2.41. *Sea $\mathcal{G} = (V, E)$ un digrafo. Decimos que $\mathcal{G}$ es débilmente conexo si al reemplazar todos sus aristas dirigidos por aristas no dirigidos obtenemos un grafo no dirigido conexo.*

2.3.2 Matriz de adyacencia e incidencia de un grafo

Quizás la forma más usual de describir un grafo, es definiendo el conjunto de vértices y el conjunto de aristas; otra forma más, sería simplemente teniendo una representación ilustrativa del grafo, es decir, tener al grafo descrito con un dibujo; pero hay otras formas más de poder describir a un grafo, las cuales son, mediante su matriz de adyacencia o su matriz de incidencia. Esta manera de representar a un grafo mediante una matriz, nos permite estudiar a los grafos desde una estructura diferente e interesante, que iremos desarrollando poco a poco.

Definición 2.42. *La matriz de adyacencia $A = (a_{ij})_{n\times n}$ de un grafo $\mathcal{G} = (V, E)$ de orden n está definida por*

$$a_{ij} := \begin{cases} 1, & \text{si } \{v_i, v_j\} \in E, \\ 0, & \text{en otro caso.} \end{cases} \tag{2.29}$$

Definición 2.43. *La matriz de incidencia $B = (b_{ij})_{n\times m}$ de un grafo $\mathcal{G} = (V, E)$ con $V = \{v_1, ..., v_n\}$ y $E = \{e_1, ..., e_m\}$ es definida sobre $\{0, 1\}$ por*

$$b_{ij} := \begin{cases} 1, & \text{si } v_i \in e_j, \\ 0, & \text{en otro caso.} \end{cases} \tag{2.30}$$

Ejemplo 2.15. *En la Figura 33 ilustramos el grafo $\mathcal{G} = (V, E)$, donde $V = \{v_1, v_2, v_3, v_4, v_5\}$ y $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$. Además, describimos la matriz de adyacencia A del grafo $\mathcal{G}$ en (2.31) y la matriz de incidencia B del grafo $\mathcal{G}$ en (2.32).*

La matriz de adyacencia A es de tamaño 5×5 pues el orden del grafo $\mathcal{G}$ es 5, además, las componentes de la matriz A son 0 ó 1, pues 0 indica que no hay una arista entre el par de vértices y 1 que si hay una arista que une a los vértices. La matriz de incidencia B es de tamaño 5×7 pues su número de renglones lo determina el número de vértices de $\mathcal{G}$ que es 5, y su número de columnas es el número de aristas de $\mathcal{G}$ que es 7, también la matriz B tiene como componentes a 0 ó 1, 0 si el vértice no es incidente con la arista y 1 si el vértice es incidente con la arista.

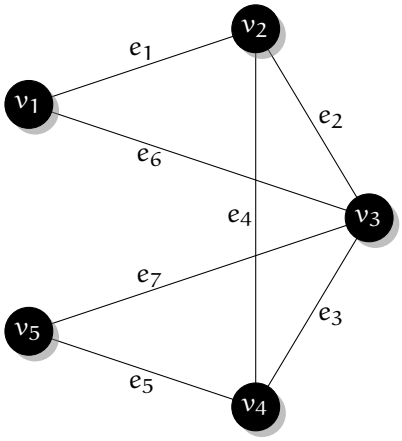


Figura 33: Grafo $\mathcal{G}$ con su respectiva matriz de adyacencia A y la matriz de incidencia B.

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \end{array} \tag{2.31}$$

$$B = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \tag{2.32}$$

2.3.3 Grafos bipartitos

Los grafos bipartitos desempeñan un papel importante dentro de los códigos lineales, y este concepto proviene de la idea general de grafos r-partitos, por lo que a continuación damos su definición.

Definición 2.44. *Sea* $r \geqslant 2$ *un entero. Un grafo* $\mathcal{G} = (V, E)$ *es llamado* r*-partito si* V *admite una partición de* r *clases tales que cada arista tiene sus extremos en diferentes clases; los vértices en la misma clase de la partición no deben ser adyacentes.*

Ejemplo 2.16. *Dado el grafo* $\mathcal{G} = (V, E)$ *con* $V = \{1, 2, 3, 4, 5, 6\}$ *y* $E = \{\{1, 3\}, \{1, 4\}, \{2, 4\}, \{5, 6\}\}$*. De acuerdo*

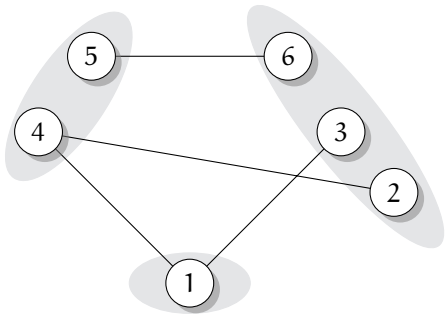


Figura 34: Grafo $\mathcal{G} = (V, E)$ 3-partito, pues el conjunto de vértices admite una partición de 3 clases.

al conjunto de aristas E *podemos observar que el conjunto de vértices* V *admite una partición de* $r = 3$ *clases,*

$V_1 = \{1\}$, $V_2 = \{2,3,6\}$, *y* $V_3 = \{4,5\}$, *pues los vértices de la misma clase no son adyacentes y los extremos de las aristas están en diferentes clases. Siendo así, conseguimos un grafo* $\mathcal{G}$ *3-partito que se ilustra en la Figura 34.*

Ahora bien, de la Definición 2.44 en particular para $r = 2$ tenemos la siguiente definición.

Definición 2.45. *Un grafo bipartito, es un grafo* $\mathcal{G} = (V, E)$ *donde el conjunto de vértices* V *puede ser dividido en dos conjuntos de vértices,* $V_1, V_2 \subseteq V$*, tales que estos dos conjuntos son disjuntos, es decir,* $V_1 \cap V_2 = \emptyset$*, y su unión es todo el conjunto* V*, esto es,* $V_1 \cup V_2 = V$*. Además, no sucede que cualquier par de vértices que estén en el mismo conjunto sean adyacentes.*

Comentario 2.18. *Podemos denotar a un grafo bipartito como* $\mathcal{G} = (V_1 \cup V_2, E)$.

Ejemplo 2.17. *Dado el grafo* $\mathcal{G} = (V, E)$ *con* $V = \{1,2,3,4,5\}$ *y* $E = \{\{1,3\},\{1,5\},\{2,3\},\{2,5\},\{4,3\},\{4,5\}\}$*. Al observar al conjunto de aristas, podemos notar que los vértices* 3 *y* 5 *nunca son adyacentes, pero estos sí se conectan a los vértices* 1, 2 *y* 4*, los cuales tampoco son adyacentes entre sí. Ahora, si consideramos los siguientes conjuntos,* $V_1 = \{1,2,4\}$ *y* $V_2 = \{3,5\}$*, que son subconjuntos de* V*, tales que son disjuntos y que su unión es todo* V*, entonces el grafo* $\mathcal{G}$ *es un grafo bipartito* $\mathcal{G} = (V_1 \cup V_2, E)$*, el cual ilustramos en la Figura 35*

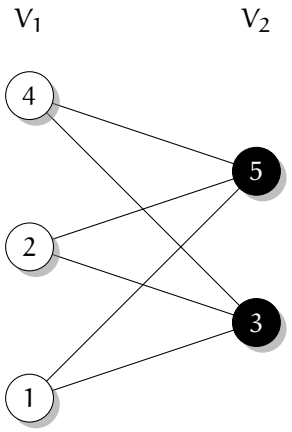


Figura 35: Grafo bipartito $\mathcal{G} = (V, E)$, ya que V admite una partición de 2 clases V_1 y V_2.

Definición 2.46. *Un grafo bipartito regular es un grafo bipartito* $\mathcal{G} = (V_1 \cup V_2, E)$*, tal que todos los vértices en* V_1 *tienen un mismo grado, denotado por* d_{V_1}*, y todos los vértices en* V_2 *también tienen un mismo grado, denotado por* d_{V_2}*, y donde los dos grados no necesariamente son iguales. Un grafo bipartito irregular es un grafo bipartito que no es regular.*

Comentario 2.19. *Un grafo bipartito regular con grados* d_{V_1} *y* d_{V_2} *de sus correspondientes conjuntos de vértices, es nombrado un grafo bipartito* (d_{V_1}, d_{V_2})*-regular.*

Observación 2.16. *Para un grafo bipartito irregular,* d_{V_1} *y* d_{V_2} *denotan los grados máximos para su conjunto de vértices correspondiente.*

A continuación se presenta un método para transformar un grafo no bipartito en un grafo bipartito a través del uso de *grafos de incidencia de arista-vértice*. De esta manera, los grafos de incidencia arista-vértice pueden ser usados para obtener un grafo bipartito a partir de grafos que no son bipartitos.

Definición 2.47. *Sea* $\mathcal{G} = (V, E)$ *un grafo con* E *un conjunto de aristas y* V *un conjunto de vértices. El grafo de incidencia arista-vértice* $\mathcal{G}' = (V', E')$ *de* $\mathcal{G}$*, es el grafo bipartito con conjunto de vértices* $V' = E \cup V$ *y conjunto de arista* $E' = \{\{e, v\} \in (E, V) : v$ *es un punto extremo de la arista* $e\}$*.*

Ejemplo 2.18. *En la Figura 36 ilustramos un grafo y su grafo de incidencia arista-vértice. Sea un grafo* $\mathcal{G} = (V, E)$*, con* $V = \{A, B, C\}$ *y* $E = \{e_1 = \{A, B\}, e_2 = \{B, C\}, e_3 = \{A, C\}\}$*. Luego, el grafo de incidencia arista-vértice de* $\mathcal{G}$ *es* $\mathcal{G}' = (V', E')$*, donde* $V' = E \cup V = \{e_1, e_2, e_3, A, B, C\}$ *y* $E' = \{\{e, v\} \in (E, V) :$ *v es un punto extremo de la arista* $e\} = \{\{e_1, A\},\{e_1, B\},\{e_2, B\},\{e_2, C\},\{e_3, A\},\{e_3, C\}\}$*.*

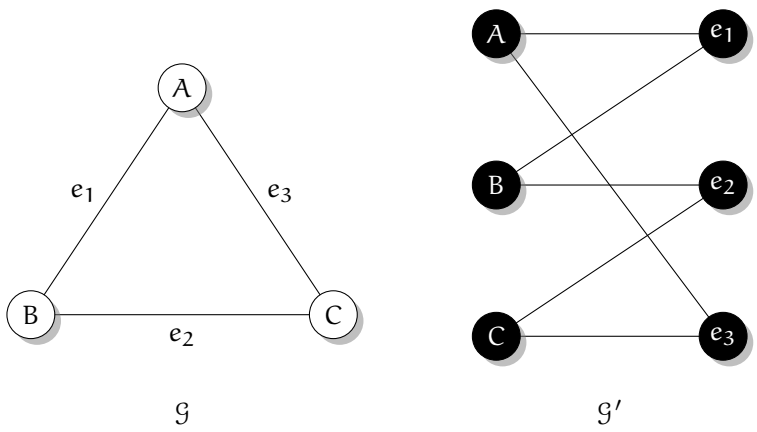


Figura 36: Al construir el grafo de incidencia arista-vértice del grafo no bipartito G, conseguimos un grafo bipartito G'.

G es un grafo 2-regular, luego, el grafo de incidencia arista-vértice G' de G, resulta ser un grafo bipartito $(2,2)$-regular.

También podemos obtener un grafo bipartito desde un grafo no bipartito a través del uso de la *cubierta doble*.

Definición 2.48. *Dado un grafo $G = (V, E)$, definamos un grafo bipartito $G' = (V', E')$, con el conjunto de vértices $V_1 \cup V_2$ tal que V_1 y V_2 son copias de V, y un vértice en V_1 y un vértice en V_2 son adyacentes solo si los vértices correspondientes en V son adyacentes en G. Y así, G' es llamado la cubierta doble de G.*

Ejemplo 2.19. *En la Figura 37 ilustramos un grafo y su cubierta doble. Sea un grafo $G = (V, E)$, con $V = \{a, b, c, d, f\}$ y $E = \{\{a, b\}, \{a, c\}, \{b, c\}, \{c, d\}, \{c, f\}, \{d, f\}\}$. Luego, la cubierta doble de G es $G' = (V', E')$,*

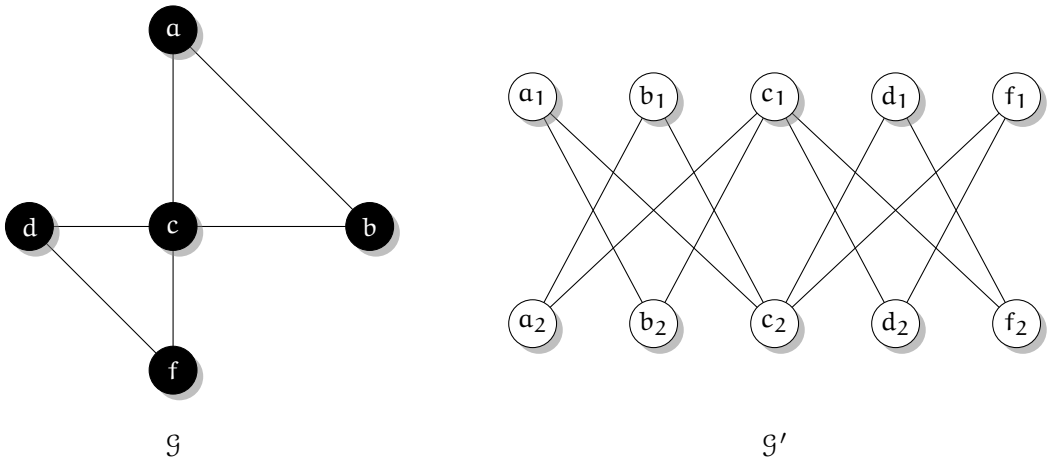


Figura 37: Al construir la cubierta doble del grafo no bipartito G, conseguimos un grafo bipartito G'.

donde $V' = V_1 \cup V_2$ con V_1 y V_2 copias de V, es decir, renombrando los vértices de V tenemos $V_1 = \{a_1, b_1, c_1, d_1, f_1\}$ y $V_2 = \{a_2, b_2, c_2, d_2, f_2\}$, así, $V' = \{a_1, b_1, c_1, d_1, f_1, a_2, b_2, c_2, d_2, f_2\}$, además, $E' = \{\{a_1, b_2\}, \{a_1, c_2\}, \{b_1, a_2\}, \{b_1, c_2\}, \{c_1, a_2\}, \{c_1, b_2\}, \{c_1, d_2\}, \{c_1, f_2\}, \{d_1, c_2\}, \{d_1, f_2\}, \{f_1, c_2\}, \{f_1, d_2\}\}$.

Notemos que G es un grafo irregular, luego, la cubierta doble G' de G, resulta ser un grafo bipartito irregular.

Comentario 2.20. *En los dos métodos anteriores, hemos notado que si comenzamos a trabajar con un grafo r-regular obtenemos un grafo bipartito (d_{V_1}, d_{V_2})-regular, y si trabajamos con un grafo irregular conseguimos un grafo bipartito irregular.*

2.4 EXPANSIÓN DE GRAFOS

En esta sección estudiamos el concepto de grafo expandido. En sí, la noción básica de expansión es que todos los conjuntos pequeños deberían tener vecindades grandes. La meta al hablar de grafos es establecer buenas propiedades de conectividad sin tener muchas aristas. Así, la construcción requeriría que se expandan por un factor constante y que tengan un número grande de vecinos con pocas aristas. Tales grafos con esta característica existen. Y es posible demostrar la existencia de una familia de expansiones lineales usando construcciones probabilísticas, pues un proceso aleatorio simple produciría uno con alta probabilidad. Sin embargo, las construcciones explícitas o deterministas son más deseables, lamentablemente estas construcciones son más díficiles de tener.

Los grafos expandidos fueron definidos por Bassalygo y Pinsker, y su primera existencia fue probada por Pinsker en los inicios de los 70's. Margulis encontró una manera para construir explícitamente expansiones lineales, con lo cual él dió la primera construcción explícita de una familia infinita de grafos expandidos.

La manera más común para probar que un grafo es una buena expansión es examinando su segundo valor propio más grande λ_G de su matriz de adyacencia. Para grafos regulares, los resultados son basados en el cálculo de λ_G. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los distintos valores propios del grafo de orden n tales que $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, y sea λ_G el segundo valor propio más grande en valor absoluto de G, [WK03]. Suponiendo que G es un grafo k-regular con n vértices. Notemos que $\lambda_G = max(\mid \lambda_2 \mid, \mid \lambda_n \mid)$ sí G no es bipartito, ya que al tener un grafo k-regular se cumple que $\lambda_1 = k$ con multiplicidad 1 y $\lambda_n \geqslant -k$, consultar Proposición 3.2 y 3.3; ahora, $\lambda_G = \mid \lambda_2 \mid$ sí G es bipartito, y además, $\lambda_1 = k$ y $\lambda_n = -k$, consultar Proposición 3.4 y Teorema 3.14.

Fue mostrado por Alon y Tanner que un grafo G es una buena expansión si y sólo si λ_1 y λ_G están separados, es decir, si el segundo valor propio más grande está distante del primer valor propio, [WK03]. Por lo tanto, para encontrar una buena expansión de un grafo G, es suficiente con checar el valor de λ_G. La siguiente cota inferior para λ_G

$$\liminf_{n\to\infty} \lambda_G \geqslant 2\sqrt{k-1},$$

es la máxima separación posible entre el primer y segundo valor propio más grande en un grafo. Esta cota fue lograda en la construcción explícita de Margulis, y Lubotzky, Phillips, y Sarnak, [SS96].

La razón por la que hablaremos de este concepto de expansión de grafos es que, antes de fijarnos en la decodificación de la palabra recibida, podemos usar la expansión de un grafo bipartito para garantizar la capacidad de corrección de error del código asociado. En sí, si el grafo bipartito es una buena expansión, entonces el código asociado será un buen código.

A partir del conocimiento de este concepto se puede acotar la distancia mínima de un código. Una ventaja es que el concepto se aplica directamente a los grafos de tamaño finito, pero la desventaja, es que la cota derivada por estos métodos es pésima y no reflejan el verdadero potencial de corrección de error observado en la práctica.

La expansión es un concepto matemático fundamental, que bien merece ser investigado a fondo propiamente, sin embargo, aquí sólo estudiaremos la definición para grafos bipartitos que más nos interesa, al igual que algunas propiedades y mostraremos algunos ejemplos.

Los textos revisados para el desarrollo de esta sección fueron [AGM87], [Alo88], [GG81], [HLW06], [Lub10] y [RU08].

2.4.1 Concepto básico de expansión de grafos bipartitos

Hay muchas variaciones de la definición de expansión de grafos, en partícular para grafos bipartitos, aunque tales definiciones son esencialmente las mismas. A continuación enunciaremos varias de estas definiciones, pero las Definiciones 2.51 y 2.52 son con las cuales trabajaremos más adelante. Definamos el concepto de expansión para grafos bipartitos regulares, comencemos con la definición cuando las dos clases de vértices tienen el mismo grado, es decir, es k-regular.

Definición 2.49. *Una* (n, k, c)*-expansión es un grafo bipartito* k*-regular* $\mathcal{G} = (I \cup O, E)$ *con* I *el conjunto de* n *vértices de entrada y* O *el conjunto de* n *vértices de salida y a lo más* $k \cdot n$ *aristas, tal que para cada subconjunto* X *de vértices de entrada,* $X \subseteq I$*, se cumple la siguiente relación,*

$$|\Gamma_X| \geqslant \left[1 + c\left(1 - \frac{|X|}{n}\right)\right]|X|, \tag{2.33}$$

donde Γ_X *es el conjunto de los vértices de salida conectados a* X*.*

Definición 2.50. *Sea* $c' > 0$*. Una* (n, k, c')*-expansión es un grafo bipartito* k*-regular* $\mathcal{G} = (I \cup O, E)$ *con* I *el conjunto de los vértices de entrada y* O *el conjunto de los vértices de salida, es decir, las aristas van de* I *a* O*, y con* $|I| = |O| = n$*, tal que, para cualquier* $X \subset I$ *con* $|X| \leqslant \frac{n}{2}$ *tenemos* $|\Gamma_X| \geqslant (1 + c')|X|$*.*

Una definición más completa.

Definición 2.51. *Una* (n, k, c)*-expansión es un grafo bipartito* k*-regular* $\mathcal{G} = (I \cup O, E)$ *sobre* I *el conjunto de vértices de entrada y* O *el conjunto de vértices de salida, donde* $|I| = |O| = n$*, y a lo más* $k \cdot n$ *aristas, tal que para cada subconjunto* X *de entradas,* $X \subseteq I$*, con* $|X| \leqslant \frac{n}{2}$*,* X *está unido a las aristas por lo menos* $|X| + c\left(1 - \frac{|X|}{n}\right)|X|$ *diferentes vértices de salida, es decir, se cumple la siguiente relación,*

$$si \quad |X| \leqslant \frac{n}{2} \quad \Longrightarrow \quad |\Gamma_X| \geqslant \left[1 + c\left(1 - \frac{|X|}{n}\right)\right]|X|, \tag{2.34}$$

donde Γ_X *denota el conjunto de todos los vecinos de los vértices conectados a* X*.*

Comentario 2.21. *Algunos de los grafos bipartitos expandidos que estudiaremos son construidos por cubiertas dobles de grafos no-bipartitos Definición 2.48, y por un conjunto de permutaciones.*

Las definiciones anteriores fueron planteadas para grafos bipartitos k-regulares, la que enunciamos a continuación es sobre grafos bipartitos (l, r)-regulares.

Definición 2.52. *Sea* $\mathcal{G} = (I \cup O, E)$ *un grafo bipartito* (l, r)*-regular con* I *el conjunto de* n *nodos de entrada de grado* l *y* O *el conjunto de* $\frac{l}{r} \cdot n$ *nodos de salida de grado* m*. Decimos que* $\mathcal{G}$ *es una* (l, r, α, γ)*-expansión si para cada subconjunto* X *de* I*,* $X \subseteq I$*, de a lo más* $\alpha \cdot n$ *nodos de entrada,* Γ_X *el conjunto de nodos de salida los cuales están conectados a* X *es al menos* $\gamma \cdot |X| \cdot l$*.*

Comentario 2.22. *La idea para expandir grafos bipartitos es clara, ver Figura 38. Un conjunto* X *de nodos de entrada tiene* $|X| \cdot l$ *aristas de salida y pueden por lo tanto están conectados a lo más* $|X| \cdot l$ *nodos de salida. Por consiguiente,* γ *representa el mínimo de tal fracción el cual se obtiene del grafo dado donde el mínimo es sobre todos los conjuntos no vacíos* X *de cardinalidad a lo más* $\alpha \cdot n$*. Si la expansión es suficientemente grande, la distancia mínima del código asociado es una fracción lineal de la longitud del bloque.*

A continuación se menciona el concepto de una familia de expansiones.

Definición 2.53. *Una familia de expansiones lineales de densidad* k *y expansión* c *es un conjunto* $\{\mathcal{G}_i\}_{i=1}^{\infty}$*, donde* $\mathcal{G}_i$ *es una* (n_i, k, c)*-expansión si* $n_i \to \infty$ *y* $\frac{n_{i+1}}{n_i} \to 1$ *cuando* $i \to \infty$*.*

Observación 2.17. *Esto es, una famila de expansiones es una familia de grafos* k*-regular (para un* k *fijo y* n *que tiene al infinito) los cuales son todos* c*-expansiones para algún* c*.*

Comentario 2.23. *Cada grafo* k*-regular finito conexo es una expansión para algún* $c > 0$ *en una forma trivial. La noción es de interés solo cuando uno considera una familia infinita de grafos. En diversas aplicaciones, se requiere una familia de* (n, k, c)*-expansiones donde* n *tiende a infinito y* k *y* c *son fijos. Usualmente (pero no siempre) se prefiere que* k *sea tan pequeño como sea posible, y es siempre deseable que* c *sea lo más grande posible.*

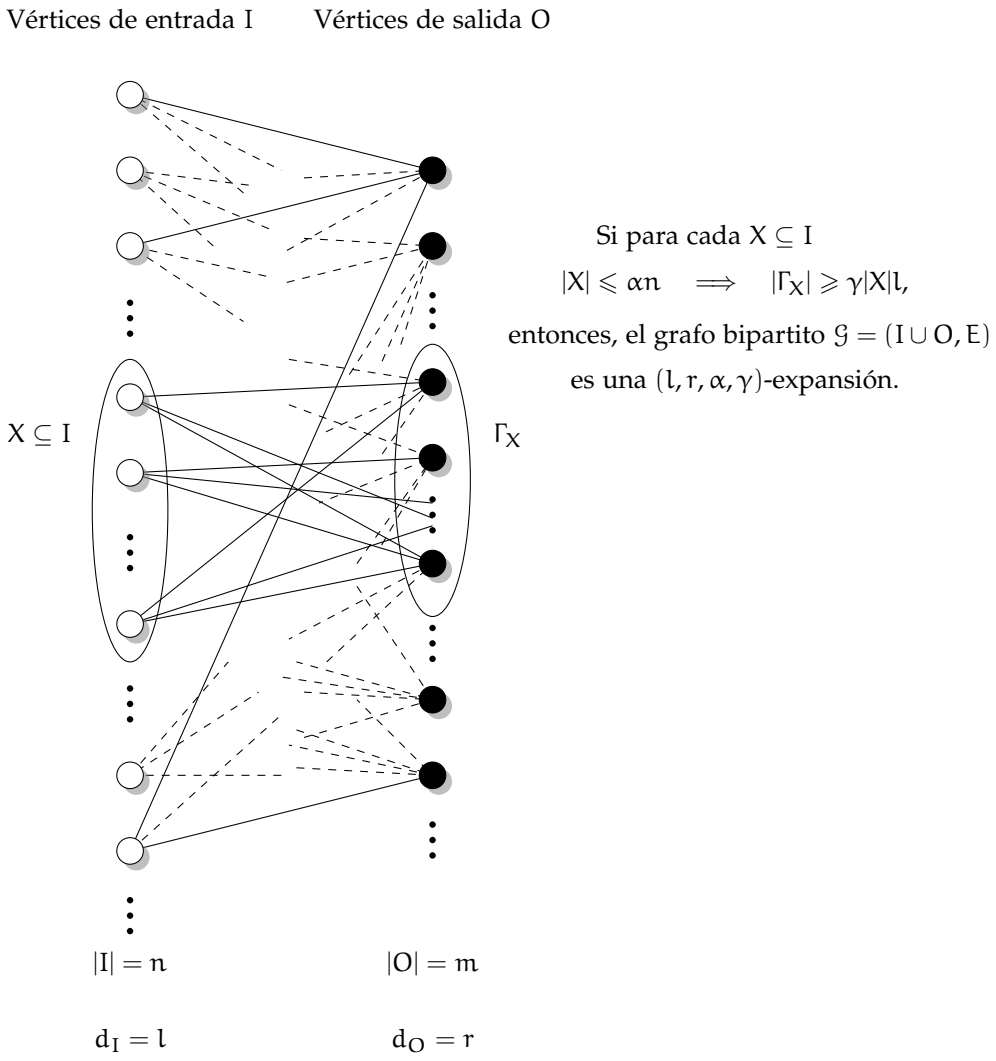


Figura 38: Concepto de expansión de un grafo bipartito.

Podemos encontrar otras definiciones de expansión de grafos en la literatura. Aunque todas son esencialmente equivalentes con algún cambio o modificación en las constantes. La idea básica es siempre que para cada subconjunto del conjunto de vértices se garantiza que se expande por alguna cantidad fija como ilustramos en la Figura 38.

2.4.2 Existencia de expansiones

Es díficil mostrar la existencia de expansiones y todavía más dar una construcción determinista. Existen dos métodos para la construcción explícita de familias de expansiones: el primero dado por Margulis, quien construyó expansiones con la ayuda de la propiedad (T) de Kazhdan de la teoría de representación de grupos de Lie semi-simples; la segunda construcción realizada por Lubotzky-Phillips-Sarnak, y Margulis, quienes usaron la conjetura de Ramanujan para este propósito, [Lub10].

El siguiente par de proposiciones son útiles para encontrar grafos expandidos; sin embargo, no abordaremos sus respectivas demostraciones pues no es el objetivo de la sección.

Proposición 2.5. *Sea* $k \geqslant 5$ *un entero y* $c = \frac{1}{2}$. *Entonces para algún grafo* k*-regular con* n *vértices se satisface que* $|\Gamma_X| \geqslant c|X|$ *para subconjuntos* X *de tamaño menor o igual a* $\frac{n}{2}$. *Para* n *grande, la mayoría de los grafos* k*-regulares con* n *vértices satisfacen lo anterior. En particular, estos grafos son expandidos.*

Demostración. Ver la demostración en [Lub10] pág. 5. □

Proposición 2.6. *Sea m un entero positivo y $V_m = \frac{\mathbb{Z}}{m\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$. Definimos un grafo sobre el conjunto V_m conectando cada $(a,b) \in V_m$ por $\sigma_1(a,b) = (a+1,b)$, $\sigma_2(a,b) = (a,b+1)$, $\sigma_3(a,b) = (a,a+b)$ y $\sigma_4(a,b) = (-b,a)$. Entonces $\{V_m\}$ es una familia de expansiones.*

Demostración. Revisar la demostración en [Lub10] pág. 32. □

2.4.3 Ejemplos de grafos expandidos

Los grafos expandidos pueden ser grafos simples, no simples, multigrafos, bipartitos, entre otros. A continuación mencionamos algunas familias distintas de grafos expandidos.

Ejemplo 2.20. *Una familia de grafos 8-regular $\mathcal{G}_m$ para cada entero m. El conjuntos de vértices es $V_m = \mathbb{Z}_m \times \mathbb{Z}_m$. Los vecinos de los vértices (a,b) son $(a+b,b)$, $(a-b,b)$, $(a,b+a)$, $(a,b-a)$, $(a+b+1,b)$, $(a-b+1,b)$, $(a,b+a+1)$, $(a,b-a+1)$, todas las operaciones son módulo m. Esta familia de grafos, debida a Margulis, es la primer construcción explícita de una familia de grafos expandidos, [HLW06] pág. 453.*

Ejemplo 2.21. *Una familia de grafos bipartitos $\mathcal{G}_n$ con n vértices donde para cada entero m, $n = m^2$ y el conjunto de vértices es $V_m = \{0,1,\ldots,m-1\} \times \{0,1,\ldots,m-1\}$. Los grafos bipartitos $\mathcal{G}_n$ son obtenidos de 5 permutaciones en V_m. Las permutaciones son: $\sigma_1(a,b) = (a,b)$, $\sigma_2(a,b) = (a,a+b)$, $\sigma_3(a,b) = (a,a+b+1)$, $\sigma_4(a,b) = (a+b,b)$, $\sigma_5(a,b) = (a+b+1,b)$, donde $+$ es módulo m, [GG81]. Esta familia de grafos bipartitos son expansiones.*

Damos algunos ejemplos concretos.

Ejemplo 2.22. *Considérese cuatro permutaciones arbitrarias $\sigma_1, \sigma_2, \sigma_3$, y σ_4 del conjunto $\{1,2,3,4,5\}$. Construimos un grafo $\mathcal{G}$ conectado con una arista del vértice $(i,0)$ al vértice $(\sigma_1(i),1)$, $(\sigma_2(i),1)$, $(\sigma_3(i),1)$ y $(\sigma_4(i),1)$, como vemos en la Tabla 10.*

	1	2	3	4	5
σ_1	3	1	4	2	5
σ_2	1	4	3	5	2
σ_3	5	1	4	2	3
σ_4	1	2	5	3	4

Tabla 10: Éstas son las 4 permutaciones con las que trabajamos.

En la Figura 39 ilustramos a $\mathcal{G} = (I \cup O, E)$ el grafo bipartito resultante que es un multigrafo pues contiene aristas múltiples. Analicemos el grafo bipartito $\mathcal{G} = (I \cup O, E)$, tiene como conjunto de vértices de entrada $I = \{1,2,3,4,5\}$, el conjunto de vértices salida es $O = \{1,2,3,4,5\}$, el conjunto de aristas es $E = \{\{1,1\},\{1,1\},\{1,3\},\{1,5\},\{2,1\},\{2,1\},\{2,2\},\{2,4\},\{3,3\},\{3,4\},\{3,4\},\{3,5\},\{4,2\},\{4,2\},\{4,3\},\{4,5\},\{5,2\},\{5,3\},\{5,4\},\{5,5\}\}$. Luego, $|I| = |O| = 5$ y los grados de los dos conjuntos de vértices son $d_I = 4$ y $d_O = 4$. Así, $\mathcal{G}$ es un grafo bipartito 4-regular con 5 vértices de entrada y de salida y a lo más $k \cdot n = 4 \cdot 5 = 20$ aristas.

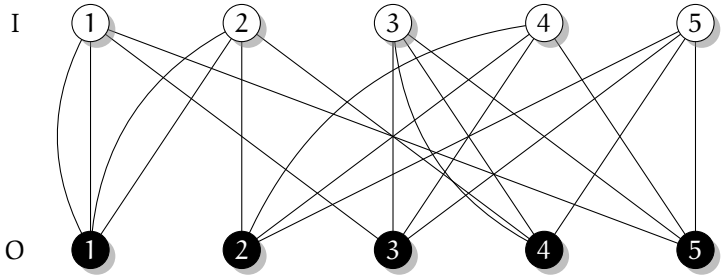


Figura 39: Grafo bipartito obtenido de 4 permutaciones.

Para verificar que es un grafo bipartito expandido tenemos que revisar que se satisface la Definición 2.51, es decir, que se cumple la siguiente relación:

$$Si \quad |X| \leqslant \frac{n}{2} \quad \Longrightarrow \quad |\Gamma_X| \geqslant \left[1 + c\left(1 - \frac{|X|}{n}\right)\right]|X|. \tag{2.35}$$

El grafo bipartito $\mathcal{G}$ satisface las condiciones de la definición. Ahora, tomemos un subconjunto X de I con $|X| \leqslant \frac{5}{2}$ y veamos que ocurre con Γ_X.

♯ *Subconjuntos*	$\lvert X\rvert$	X		Γ_X	$\lvert\Gamma_X\rvert$
$\binom{5}{1} = 5$	1	$\{1\}$	$\to$	$\{1,3,5\}$	3
	1	$\{2\}$	$\to$	$\{1,2,4\}$	3
	1	$\{3\}$	$\to$	$\{3,4,5\}$	3
	1	$\{4\}$	$\to$	$\{2,3,5\}$	3
	1	$\{5\}$	$\to$	$\{2,3,4,5\}$	4

♯ *Subconjuntos*	$\lvert X\rvert$	X		Γ_X	$\lvert\Gamma_X\rvert$
$\binom{5}{2} = 10$	2	$\{1,2\}$	$\to$	$\{1,2,3,4,5\}$	5
	2	$\{1,3\}$	$\to$	$\{1,3,4,5\}$	4
	2	$\{1,4\}$	$\to$	$\{1,2,3,5\}$	4
	2	$\{1,5\}$	$\to$	$\{1,2,3,4,5\}$	5
	2	$\{2,3\}$	$\to$	$\{1,2,3,4,5\}$	5
	2	$\{2,4\}$	$\to$	$\{1,2,3,4,5\}$	5
	2	$\{2,5\}$	$\to$	$\{1,2,3,4,5\}$	5
	2	$\{3,4\}$	$\to$	$\{2,3,4,5\}$	4
	2	$\{3,5\}$	$\to$	$\{2,3,4,5\}$	4
	2	$\{4,5\}$	$\to$	$\{2,3,4,5\}$	4

Ahora, tenemos que hallar el valor de c que cumpla con la relación (2.35), por lo que procedemos de la manera siguiente:

- *para $|X| = 1$ tenemos que $|\Gamma_X| = 3, 4$, entonces,*
  $3 \geqslant \left[1 + c\left(1 - \frac{1}{5}\right)\right] 1 \Rightarrow 2 \geqslant \frac{4}{5}c \Rightarrow c \leqslant \frac{5}{2}$,
  $4 \geqslant \left[1 + c\left(1 - \frac{1}{5}\right)\right] 1 \Rightarrow 3 \geqslant \frac{4}{5}c \Rightarrow c \leqslant \frac{15}{4}$;
- *para $|X| = 2$ tenemos que $|\Gamma_X| = 4, 5$, entonces,*
  $4 \geqslant \left[1 + c\left(1 - \frac{2}{5}\right)\right] 2 \Rightarrow 2 \geqslant \frac{6}{5}c \Rightarrow c \leqslant \frac{5}{3}$,
  $5 \geqslant \left[1 + c\left(1 - \frac{2}{5}\right)\right] 2 \Rightarrow 3 \geqslant \frac{6}{5}c \Rightarrow c \leqslant \frac{5}{2}$;

luego, tomando a $c = \min\{\frac{5}{2}, \frac{15}{4}, \frac{5}{3}\}$, tenemos que el factor constante que satisface la relación es $c = \frac{5}{3}$. Por lo tanto el grafo bipartito 4-regular $\mathcal{G}$ es una $(5, 4, \frac{5}{3})$-expansión.

Ejemplo 2.23. *Sean $m = 2$ y $V_2 = \{0,1\} \times \{0,1\} = \{00, 10, 01, 11\}$, luego, $n = m^2 = 4$ es el número de vértices para el grafo el cual aún no es un grafo bipartito. Ahora sean $\sigma_1 = (a, b)$, $\sigma_2 = (a, a + b)$, $\sigma_3 = (a, a + b + 1)$, $\sigma_4 = (-b, a)$ y $\sigma_5 = (a + b + 1, b)$, las 5 permutaciones sobre V_2 usando aritmética módulo 2, la Tabla 11 muestra los resultados de las permutaciones:*

	00	10	01	11
σ_1	00	10	01	11
σ_2	00	11	01	01
σ_3	01	10	00	11
σ_4	00	10	11	10
σ_5	10	00	01	11

Tabla 11: Resultados de las permutaciones.

Ahora, con ayuda de la rejilla de puntos y las permutaciones, veamos que los vecinos del punto 10 *son* 10, 11, 10, 10, *y* 00, *con los cuales forma aristas entre ellas lazos y enlaces como ilustramos en la Figura 40.*

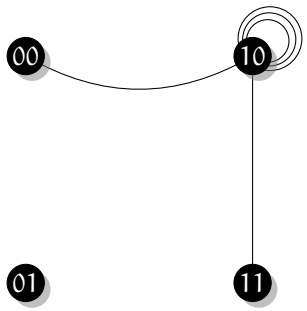


Figura 40: Rejilla de puntos de las permutaciones.

Al continuar realizando lo mismo para los puntos restantes contruimos el grafo mostrado en la Figura 41 que es un multigrafo.

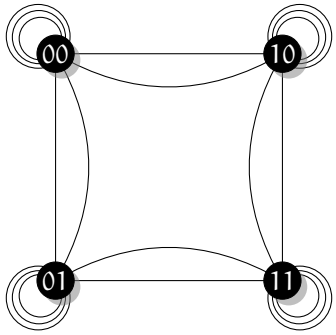


Figura 41: Grafo resultante.

Luego, renombremos a cada vértice de una manera distinta, para $00 \to 0$, $10 \to 1$, $01 \to 2$ *y* $11 \to 3$, *así, el conjunto de vértices es* $V_2 = \{0, 1, 2, 3\}$, *y de esta manera la tabla anterior ahora sería la Tabla 12:*

	0	1	2	3
σ_1	0	1	2	3
σ_2	0	3	2	1
σ_3	2	1	0	3
σ_4	0	1	3	2
σ_5	1	0	2	3

Tabla 12: Renombrando los vértices de la Tabla 11.

Y usando el concepto de la cubierta doble dado en la Definición 2.48, construimos el grafo bipartito $\mathcal{G}_2$ *que ilustramos en la Figura 42. Analicemos el grafo bipartito* $\mathcal{G}_2 = (I \cup O, E)$, *tiene como conjunto de vértices de entrada* $I = V_2$, *el conjunto de vértices salida es* $O = V_2$, *el conjunto de aristas es* $E = \{\{0,0\},\{0,0\},\{0,0\},\{0,1\},\{0,1\},\{0,2\},\{0,2\},\{1,0\},\{1,0\},\{1,1\},\{1,1\},\{1,1\},\{1,3\},\{1,3\},\{2,0\},\{2,0\},\{2,2\},\{2,2\},\{2,2\},\{2,3\},\{2,3\},\{3,1\},\{3,1\},\{3,2\},\{3,2\},\{3,3\},\{3,3\},\{3,3\}\}$. *Luego,* $|I| = |O| = 4$ *y los grados de los dos conjuntos de vértices son* $d_I = 7$ *y* $d_O = 7$. *Además, observamos que sigue siendo un multigrafo. Así,* $\mathcal{G}_2$ *es un grafo bipartito 7-regular con* 4 *vértices de entrada y de salida y a lo más* $k \cdot n = 7 \cdot 4 = 28$ *aristas.*

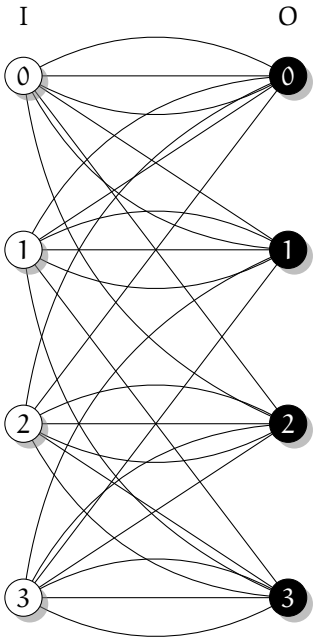


Figura 42: Grafo bipartito obtenido de la cubierta doble.

Para verificar que es un grafo bipartito expandido tenemos que revisar que se satisface la Definición 2.51, es decir, que se cumple la relación (2.35) enunciada en el ejemplo anterior. Entonces, considerando que el grafo bipartito $\mathcal{G}_2$ cumple con las condiciones de la definición, tomemos los subconjuntos X de I tales que $|X| \leqslant \frac{4}{2} = 2$ y veamos que ocurre con Γ_X.

♯ *Subconjuntos*

$$\binom{4}{1} = 4$$

$\lvert X\rvert$	X		Γ_X	$\lvert\Gamma_X\rvert$
1	$\{0\}$	$\to$	$\{0,1,2\}$	3
1	$\{1\}$	$\to$	$\{0,1,3\}$	3
1	$\{2\}$	$\to$	$\{0,2,3\}$	3
1	$\{3\}$	$\to$	$\{1,2,3\}$	3

♯ *Subconjuntos*

$$\binom{4}{2} = 6$$

$\lvert X\rvert$	X		Γ_X	$\lvert\Gamma_X\rvert$
2	$\{0,1\}$	$\to$	$\{0,1,2,3\}$	4
2	$\{0,2\}$	$\to$	$\{0,1,2,3\}$	4
2	$\{0,3\}$	$\to$	$\{0,1,2,3\}$	4
2	$\{1,2\}$	$\to$	$\{0,1,2,3\}$	4
2	$\{1,3\}$	$\to$	$\{0,1,2,3\}$	4
2	$\{2,3\}$	$\to$	$\{0,1,2,3\}$	4

Ahora, tenemos que hallar el valor de c que cumpla con la relación (2.35), por lo que procedemos de la manera siguiente:

- *para $|X| = 1$ tenemos que $|\Gamma_X| = 3$, así,*
  $3 \geqslant \left[1 + c\left(1 - \frac{1}{4}\right)\right] 1 \Rightarrow 2 \geqslant \frac{3}{4}c \Rightarrow c \leqslant \frac{8}{3}$;
- *para $|X| = 2$ tenemos que $|\Gamma_X| = 4$, así,*
  $4 \geqslant \left[1 + c\left(1 - \frac{2}{4}\right)\right] 2 \Rightarrow 2 \geqslant c \Rightarrow c \leqslant 2$;

entonces tomemos a $c = \min\{\frac{8}{3}, 2\}$. Así, el factor constante que satisface la relación es $c = 2$. Por lo tanto el grafo bipartito 7-regular $\mathcal{G}$ es una $(4, 7, 2)$-expansión.

2.5 RELACIÓN ENTRE CÓDIGOS LINEALES Y GRAFOS BIPARTITOS

La interpretación teórica que se da entre los códigos lineales y los grafos bipartitos es importante, no sólo para la construcción de códigos sino también para el diseño de algoritmos eficientes de decodificación, ya que los grafos bipartitos representan o están asociados a algún código lineal; y sucede que, el proceso de codificación y decodificación del código lineal puede basarse en esta representación del grafo bipartito asociado.

Entonces, existe una relación interesante entre los códigos lineales y los grafos bipartitos. Consideremos la matriz de chequeo de paridad H de un código detector-corrector de error $\mathcal{C}$, y sean los conjuntos V_1 y V_2 que denotan las columnas y renglones de la matriz H, respectivamente. Entonces, podemos definir un grafo bipartito $\mathcal{G} = (V_1 \cup V_2, E)$, tal que $u_j \in V_1$ y $v_i \in V_2$ forman una arista $\{u_j, v_i\} \in E$, si la componente de la matriz de chequeo de paridad H del código $\mathcal{C}$, en la columna j y el renglón i es 1, y en caso de que la componente sea 0 los vértices no forman una arista. Inversamente, dado un grafo bipartito $\mathcal{G} = (V_1 \cup V_2, E)$, podemos definir una matriz binaria cuyas columnas están relacionadas con los elementos de V_1 y los renglones con los elementos de V_2, y cuya componente de la matriz en la columna j y el renglón i es 1 si $\{u_j, v_i\} \in E$, en caso contrario será 0, esta matriz binaria es la matriz de chequeo de paridad H que define un código lineal $\mathcal{C}$, [WK03].

Por consiguiente, uno puede definir un código lineal desde un grafo bipartito y viceversa. Esto es, considérese una matriz de chequeo de paridad $H = (h_{ij})$ de un código lineal $\mathcal{C}$ y un grafo bipartito $\mathcal{G} = (V_1 \cup V_2, E)$, donde $\{u_j, v_i\} \in E$ con $u_j \in V_1$ y $v_i \in V_2$ (V_1 representa la clase de vértices columna y V_2 la clase de vértices reglón en la matriz H) si y sólo si $h_{ij} \neq 0$, entonces la matriz de adyacencia A' del grafo bipartito $\mathcal{G}$ es:

$$A' = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix}.$$

Con la siguiente proposición se puede asegurar que si tenemos una matriz de adyacencia A de un grafo $\mathcal{G}$, y existe una permutación tal que puede llevar a la matriz A a la forma $A' = \left(\begin{smallmatrix} 0 & H \\ H^T & 0 \end{smallmatrix}\right)$, eso implicará que $\mathcal{G}$ es un grafo bipartito.

Proposición 2.7. *([ADH98], pág. 16) Sea $\mathcal{G}$ un grafo con vertices $v_1, v_2, \ldots, v_n$ y $A = (a_{ij})$ una matriz de adyacencia. Entonces $\mathcal{G}$ es un grafo bipartito si y sólo si existe una permutación π del conjunto $\{1, 2, \ldots, n\}$ así que la matriz $A' = (a_{\pi(i)\pi(j)})$ tiene la siguiente forma:*

$$A' = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix}, \tag{2.36}$$

donde H^T es la transpuesta de H.

Comentario 2.24. *A la representación de los códigos lineales mediante grafos bipartitos se le llama grafo de Tanner. Además, existe una relación aún más estrecha entre la matriz de chequeo de paridad H del código lineal $\mathcal{C}$ y el grafo bipartito $\mathcal{G}$, pues la matriz de chequeo de paridad H del código lineal $\mathcal{C}$ puede verse como una matriz de adyacencia A del grafo bipartito $\mathcal{G}$ asociado.*

Veamos algunos ejemplos sobre esta interesante relación.

Ejemplo 2.24. *Sea H la matriz de chequeo de paridad de un código lineal $\mathcal{C}$ de tamaño 3×4, con $V_1 = \{u_1, u_2, u_3, u_4\}$ el conjunto que denota a las columnas y $V_2 = \{v_1, v_2, v_3\}$ el conjunto que denota los renglones, dada por:*

$$H = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{cccc} u_1 & u_2 & u_3 & u_4 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Así, teniendo la matriz H *es posible construir el grafo bipartito* $\mathcal{G} = (V_1 \cup V_2, E)$ *conocido como grafo de Tanner, ilustrado en la Figura 43, recordar que por cada* 1 *en* H *se tiene una arista relacionada con la columna y el renglón respectivos:*

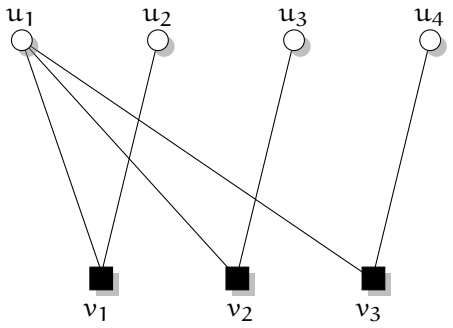


Figura 43: A partir de la matriz de chequeo de paridad H de un código lineal $\mathcal{C}$, construimos el grafo bipartito $\mathcal{G}$ y su matriz de adyacencia A. Este grafo también es conocido como grafo de Tanner debido a que representa a un código lineal $\mathcal{C}$.

Y también, obtenemos la matriz de adyacencia A *del grafo bipartito* $\mathcal{G}$*:*

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} v_1 & v_2 & v_3 & u_1 & u_2 & u_3 & u_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Ejemplo 2.25. *Ahora, definamos el grafo bipartito* $\mathcal{G} = (V_1 \cup V_2, E)$*, donde* $V_1 = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7\}$*,* $V_2 = \{r_1, r_2, r_3\}$*, y* $E = \{\{t_1, r_1\}, \{t_1, r_3\}, \{t_2, r_1\}, \{t_2, r_2\}, \{t_3, r_1\}, \{t_3, r_2\}, \{t_3, r_3\}, \{t_4, r_2\}, \{t_4, r_3\}, \{t_5, r_1\}, \{t_6, r_2\}, \{t_7, r_3\}\}$*, Figura 44.*

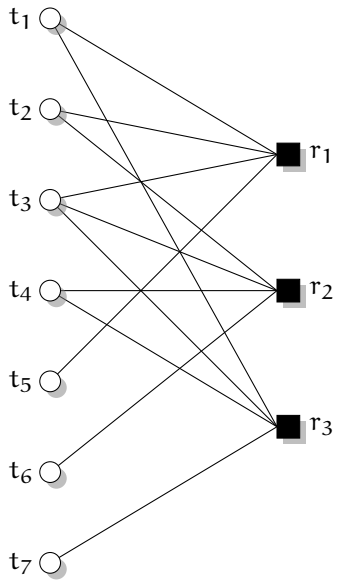


Figura 44: A partir del grafo bipartito $\mathcal{G}$, construimos la matriz de chequeo de paridad H de un código lineal $\mathcal{C}$, y así, la matriz de adyacencia A del grafo bipartito $\mathcal{G}$. Además, $\mathcal{G}$ es un grafo de Tanner y H puede verse como una matriz de adyacencia de $\mathcal{G}$.

Entonces, dado el grafo bipartito $\mathcal{G}$ es posible obtener la matriz de chequeo de paridad H de un código lineal $\mathcal{C}$, recordar que V_1 define las columnas de H y V_2 define los renglones, así:

$$H = \begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \end{array} \begin{array}{c} \begin{array}{ccccccc} t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array}, \tag{2.37}$$

en sí, la matriz de chequeo de paridad H puede verse como una matriz de adyacencia de $\mathcal{G}$, sin embargo, la matriz de adyacencia A del grafo bipartito $\mathcal{G}$ es:

$$A = \begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \end{array} \begin{array}{c} \begin{array}{cccccccccc} r_1 & r_2 & r_3 & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Antes de concluir el ejemplo, notemos que la matriz H (2.37) es la matriz de chequeo de paridad del $[7,4]$-código de Hamming obtenida desde el capítulo 1. En sí, el grafo bipartito dado en la Figura 44 tiene una forma distinta del grafo bipartito que obtuvimos en la Figura 22, pero en realidad sólo separamos las clases de vértices. Entonces, el código lineal que representa este grafo bipartito $\mathcal{G}$, Figura 44, es el $[7,4,3]$-código de Hamming.

3 ALGUNOS RESULTADOS DE ÁLGEBRA LINEAL Y DEL ESPECTRO DE UN GRAFO

Las matemásticas poseen no sólo la verdad,
sino cierta belleza suprema.

– Bertrand Arthur William Russell –

En este capítulo hablaremos de los valores propios y vectores propios de una matriz, de la diagonalización y ortogonalización, enunciaremos algunos teoremas que involucran estos conceptos para después enunciar y demostrar el Teorema de Descomposición Espectral, y finalmente todos los conceptos serán de utilidad para poder definir el espectro de un grafo, ya que el espectro de un grafo proporciona mucha información acerca de éste y dicha información será usada posteriormente. Algunos ejemplos mostrados serán realizados con ayuda del software *wxMaxima 16.04.2* [VLIK15].

La teoría dada en este capítulo está enunciada en términos de matrices debido a que las matrices describen un grafo y algunos grafos están asociados a un código, entonces al trabajar los resultados directamente con la matriz nos hará más fácil revisar los resultados para códigos. Sin embargo, estos conceptos y resultados pueden encontrarse enunciados para un operador lineal, pero esto no debería causar ningún problema pues recordemos que un operador lineal siempre tiene asociado una matriz, y es indistingible estar trabajando con un operador lineal o una matriz y viceversa, siempre y cuando estemos trabajando sobre espacios vectoriales de dimensión finita.

3.1 ALGUNOS CONCEPTOS Y RESULTADOS DE ÁLGEBRA LINEAL

Comenzaremos con las definiciones de valor propio, vector propio, polinomio característico, y espacio propio. Enunciaremos algunos resultados esenciales que usaremos y después definiremos la diagonalización y ortogonalización.

Los textos revisados para el desarrollo de esta sección fueron [BC09], [HK71] y [P0011].

3.1.1 Valores propios y vectores propios

Definición 3.1. *Sea $A = (a_{ij})$ una matriz de tamaño n. Sea λ un número real o complejo. Entonces λ es un valor propio de A cuando existe un vector diferente de cero v en $\mathbb{R}^n$ o $\mathbb{C}^n$ tal que $Av = \lambda v$. El vector no cero v es llamado el vector propio de A correspondiente al valor propio λ.*

Comentario 3.1. *Si el valor propio λ es un número real, entonces el vector propio v es un vector con componentes reales. Sin embargo, si λ es un número complejo, que puede ocurrir aún si la matriz A es de componentes reales, el vector propio v puede ser un vector con componentes complejas.*

Definición 3.2. *Si $A = (a_{ij})$ es una matriz de tamaño n entonces $P_A(t) = \det(tI_n - A)$ es un polinomio en t de grado n y es llamado el polinomio característico de A.*

Observación 3.1. *Recordemos lo siguiente,*

$$
\begin{aligned}
\lambda \text{ es una raíz de } P_A(t) &\Leftrightarrow P_A(\lambda) = 0 \\
&\Leftrightarrow \det(\lambda I - A) = 0 \\
&\Leftrightarrow \lambda I - A \text{ no es invertible} \\
&\Leftrightarrow \operatorname{Ker}(\lambda I - A) \neq \{0\} \\
&\Leftrightarrow \exists v \in V - \{0\} : (\lambda I - A)(v) = 0 \\
&\Leftrightarrow \lambda I v - Av = 0 \\
&\Leftrightarrow Av = \lambda v \\
&\Leftrightarrow \lambda \text{ es un valor propio de } A.
\end{aligned}
$$

Comentario 3.2. *Como $P_A(t)$ es un polinomio de grado n, $P_A(t)$ tiene n raíces no necesariamente distintas, entonces A tiene n valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$. Así, el conjunto de los n valores propios conforma el espectro de A.*

Definición 3.3. *Sea $A = (a_{ij})$ una matriz de tamaño n y sea λ un valor propio de A. La multiplicidad de λ es su multiplicidad como una raíz del polinomio característico $P_A(t)$. Por lo tanto la multiplicidad de un valor propio de A es un entero entre 1 y n. El espacio propio de A correspondiente a λ está definido como $V_\lambda = \{x \in \mathbb{C}^n : (\lambda I_n - A)x = 0\}$. El espacio propio de A correspondiente a λ es el espacio nulo de la matriz $\lambda I_n - A$, y así es un subespacio de $\mathbb{C}^n$. El espacio propio consiste del vector cero y todos los vectores propios de A correspondientes a λ, así, tiene dimensión de al menos 1.*

Comentario 3.3. *Si todos los n valores propios de A son números reales, entonces los espacios propios son subespacios de $\mathbb{R}^n$. Pero si A tiene un valor propio complejo, entonces los espacios propios se consideran como subespacios de $\mathbb{C}^n$. Además, la multiplicidad geométrica de λ es la dimensión del espacio propio V_λ.*

Ahora, enunciaremos algunos teoremas que nos proporcionan propiedades y características interesantes que nos serán de utilidad para hablar sobre el espectro de un grafo.

Teorema 3.1. *Sea $A = (a_{ij})$ una matriz simétrica y de valores reales de tamaño $n \times n$. Entonces cada uno de los n valores propios de A es un número real.*

Demostración. Supóngase que λ es un valor propio de A con su correspondiente vector propio v (un vector no cero en $\mathbb{C}^n$). Entonces $Av = \lambda v$ y, al tomar los conjugados complejos, se tiene $\overline{Av} = \overline{\lambda v}$. Luego, $A\bar{v} = \bar{A}\bar{v} = \overline{Av} = \overline{\lambda v} = \bar{\lambda}\bar{v}$ pues A es una matriz con entradas reales. Al tomar las transpuestas y usar el hecho de que A es simétrica, se tiene $\bar{v}^T A = \bar{v}^T A^T = (A\bar{v})^T = (\bar{\lambda}\bar{v})^T = \bar{\lambda}\bar{v}^T$. Por lo tanto, $\lambda(\bar{v}^T v) = \bar{v}^T(\lambda v) = \bar{v}^T(Av) = (\bar{v}^T A)v = (\bar{\lambda}\bar{v}^T)v = \bar{\lambda}(\bar{v}^T v)$, eso implica que $(\lambda - \bar{\lambda})(\bar{v}^T v) = 0$.

Ahora, si $v = \begin{pmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{pmatrix}$ entonces $\bar{v} = \begin{pmatrix} a_1 - b_1 i \\ \vdots \\ a_n - b_n i \end{pmatrix}$, de modo que $\bar{v}^T v = (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \neq 0$ pues $v \neq 0$ porque es un vector propio. Se concluye que $\lambda - \bar{\lambda} = 0$ eso implica que $\lambda = \bar{\lambda}$. En consecuencia, λ es un número real. □

Teorema 3.2. *([CDS80], pág. 17) La multiplicidad geométrica y algebraica de un valor propio de una matriz simétrica son iguales.*

Lema 3.1. *([CDS80], pág. 61) Si $A = (a_{ij})$ es una matriz de tamaño $m \times n$ y $P_X(t)$ denota el polinomio característico de una matriz cuadrada X, entonces*

$$t^n P_{AA^T}(t) = t^m P_{A^T A}(t). \tag{3.1}$$

Lema 3.2. *Sea $A = (a_{ij})$ una matriz de tamaño $n \times n$. Si λ es un valor propio de A^2 entonces $\pm\sqrt{\lambda}$ es un valor propio de A.*

Demostración. Sea λ un valor propio de A^2, es decir, $A^2 v = \lambda v$. Considerando el polinomio característico de A, tenemos,

$$\begin{aligned} P_{A^2}(t) &= \det(tI_n - A^2) \\ &= |tI_n - A^2| \\ &= |(\sqrt{t}I_n)^2 - A^2| \\ &= |(\sqrt{t}I_n - A)(\sqrt{t}I_n + A)| \\ &= |\sqrt{t}I_n - A| \cdot |\sqrt{t}I_n + A| \\ &= |\sqrt{t}I_n - A| \cdot |A - (-\sqrt{t})I_n| \\ &= P_A(\sqrt{t}) \cdot P_A(-\sqrt{t}) \end{aligned}$$

Entonces, $P_{A^2}(t) = P_A(\sqrt{t}) \cdot P_A(-\sqrt{t})$. Luego, si λ es un valor propio de A^2, entonces λ es una raíz de P_{A^2}, es decir, $P_{A^2}(\lambda) = 0$, además, $P_{A^2}(\lambda) = P_A(\sqrt{\lambda}) \cdot P_A(-\sqrt{\lambda}) = 0$, de aquí, $P_A(\sqrt{\lambda}) \cdot P_A(-\sqrt{\lambda}) = 0$, eso implica que $P_A(\sqrt{\lambda}) = 0$ ó $P_A(-\sqrt{\lambda}) = 0$. De esto último se cumple que $\pm\sqrt{\lambda}$ es una raíz de $P_A(t)$ y así es un valor propio de A. □

Lema 3.3. *Sea* $B = (b_{ij})$ *una matriz arbitraria de entradas reales de tamaño* $m \times n$. *Si* v *es un vector propio de* B^TB *con valor propio* $\lambda \neq 0$, *entonces* $v' = Bv$ *es un vector propio de* BB^T *con el mismo valor propio* λ.

Demostración. Sea v un vector propio de B^TB correspondiente al valor propio $\lambda \neq 0$, entonces $B^TBv = \lambda v$. Ahora, si $v' = Bv$, tenemos que $BB^Tv' = BB^T(Bv) = B(B^TBv) = B(\lambda v) = \lambda Bv = \lambda v'$. Por lo tanto, v' es un vector propio de BB^T con el mismo valor propio λ. □

Comentario 3.4. *Este resultado indica, que los valores propios distintos de cero de* B^TB *son los mismos valores propios de* BB^T.

Definición 3.4. *Una matriz* $A = (a_{ij})$ *es positiva si* $A > 0$, *es decir, si todas sus entradas son positivas. Y* A *es no negativa si* $A \geqslant 0$, *es decir, si todas sus entradas son mayores o iguales a cero.*

Definición 3.5. *Una matriz* $A = (a_{ij})$ *es primitiva si para algún* k *tenemos que* $A^k > 0$.

Definición 3.6. *Una matriz* $A = (a_{ij})$ *de tamaño* $n \times n$ *es reducible si, dada alguna permutación de los renglones y la misma permutación de las columnas,* A *puede escribirse en forma de bloque como*

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{3.2}$$

donde B *y* D *son cuadradas. O bien, de manera equivalente,* A *es reducible si hay alguna matriz de permutación* P *tal que*

$$PAP^T = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \tag{3.3}$$

Y si A *no es reducible será llamada irreducible.*

El siguiente teorema es conocido como el Teorema de Perron, que indica propiedades esenciales de un cierto valor propio λ_0 de A. Enunciaremos el teorema sin demostración debido a que no es el objetivo de este trabajo. Pero, dejamos la referencia en donde el lector interesado podrá revisar la demostración.

Teorema 3.3. *Sea* $A = (a_{ij})$ *una matriz positiva de tamaño* $n \times n$. *Entonces* A *tiene un valor propio real* λ_0 *con las siguientes propiedades:*

i) $\lambda_0 > 0$.

ii) λ_0 *tiene un vector propio positivo correspondiente.*

iii) Si λ *es cualquier otro valor propio de* A, *entonces* $|\lambda| \leqslant \lambda_0$.

Demostración. Podemos revisar la demostración en [Poo11] pág. 344. □

El siguiente teorema proporciona más propiedades de un cierto valor propio λ_0 de A. Este teorema es conocido como el Teorema de Perron-Frobenius, el cual enunciamos a continuación.

Teorema 3.4. *Sea* $A = (a_{ij})$ *una matriz de tamaño* $n \times n$ *no negativa irreducible. Entonces* A *tiene un valor propio real* λ_0 *con las siguientes propiedades:*

i) $\lambda_0 > 0$.

ii) λ_0 *tiene un correspondiente vector propio positivo.*

iii) λ_0 *tiene multiplicidad algebraica y multiplicidad geométrica igual a* 1.

iv) Para cada valor propio λ *de* A, *tenemos* $|\lambda| \leqslant \lambda_0$. *Si* A *es primitiva, entonces esta desigualdad es estricta.*

v) Si λ *es un valor propio de* A *tal que* $|\lambda| = \lambda_0$, *entonces* λ *es una raíz compleja de la ecuación* $\lambda^n - \lambda_0^n = 0$.

Comentario 3.5. *Podemos encontrar diversas versiones del Teorema de Perron-Frobenius, así como de su demostración, por ejemplo, consultar la referencia [BH12] pág. 22.*

Comentario 3.6. *Los Teoremas 3.3 y 3.4 hablan de un valor propio en especial, λ_0 que es llamado radio espectral de la matriz A.*

Definición 3.7. *Sea $A = (a_{ij})$ una matriz de tamaño $n \times n$ con sus valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$. El radio espectral λ_0 de A es el máximo de los valores absolutos de sus valores propios,*

$$\lambda_0 = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}. \tag{3.4}$$

3.1.2 Diagonalización

En esta subsección mencionamos algunos resultados conocidos de álgebra lineal sobre la diagonalización de una matriz A, que posteriormente nos serán útiles.

Definición 3.8. *Sean A y B matrices de tamaño $n \times n$. Se dice que A es semejante a B si existe una matriz invertible C de tamaño $n \times n$ tal que $C^{-1}AC = B$. Si A es semejante a B, se escribe $A \sim B$.*

Definición 3.9. *Una matriz A de tamaño $n \times n$ es diagonalizable si existe una matriz diagonal D tal que A sea semejante a D, esto es, si existe una matriz invertible C de tamaño $n \times n$ tal que $C^{-1}AC = D$.*

Teorema 3.5. *Sea A una matriz de tamaño $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_m$ distintos valores propios de A con sus correspondientes vectores propios $v_1, v_2, \ldots, v_m$. Entonces $v_1, v_2, \ldots, v_m$ son linealmente independientes.*

Demostración. La demostración es indirecta. Se supondrá que $v_1, v_2, \ldots, v_m$ son linealmente dependientes y se demostrará que esta suposición conduce a una contradicción.

Si $v_1, v_2 \ldots, v_m$ son linealmente dependientes, entonces uno de dichos vectores debe poder expresarse como una combinación lineal de los anteriores. Sea v_{k+1} el primero de los vectores v_i que pueden expresarse de esa forma. En otras palabras, $v_1, v_2, \ldots, v_k$ son linealmente independientes, luego existen escalares $c_1, c_2, \ldots, c_k$ tales que

$$v_{k+1} = c_1v_1 + c_2v_2 + \cdots + c_kv_k. \tag{3.5}$$

Al multiplicar por la izquierda en ambos lados de la ecuación (3.5) por A y con el hecho de que $Av_i = \lambda_i v_i$ para cada i, se tiene

$$\begin{aligned} \lambda_{k+1}v_{k+1} = Av_{k+1} &= A(c_1v_1 + c_2v_2 + \cdots + c_kv_k) \\ &= c_1Av_1 + c_2Av_2 + \cdots + c_kAv_k \\ &= c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots + c_k\lambda_kv_k \end{aligned} \tag{3.6}$$

Ahora multiplicando ambos lados de la ecuación (3.5) por λ_{k+1} obtenemos

$$\lambda_{k+}v_{k+1} = c_1\lambda_{k+1}v_1 + c_2\lambda_{k+1}v_2 + \cdots + c_k\lambda_{k+1}v_k \tag{3.7}$$

Cuando se resta la ecuación (3.6) de la ecuación (3.7), se obtiene

$$0 = c_1(\lambda_1 - \lambda_{k+1})v_1 + c_2(\lambda_2 - \lambda_{k+1})v_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})v_k.$$

Dado que los valores propios λ_i son todos diferentes, los términos entre paréntesis $(\lambda_i - \lambda_{k+1})$, $i = 1, \ldots, k$, son todos distintos de cero. Por tanto, $c_1 = c_2 = \cdots = c_k = 0$. Esto implica que $v_{k+1} = c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0v_1 + 0v_2 + \cdots + 0v_k$ lo que es imposible, pues el vector propio v_{k+1} no puede ser cero. En consecuencia, se tiene una contradicción, lo cual significa que la suposición de que $v_1, v_2, \cdots, v_m$ son linealmente dependientes es falsa. Por lo tanto, $v_1, v_2, \cdots, v_m$ deben ser linealmente independientes. □

Teorema 3.6. *Sea A una matriz de tamaño $n \times n$. Entonces A es diagonalizable si y sólo si A tiene n vectores propios linealmente independientes, es decir, es una base de $\mathbb{C}^n$ (o $\mathbb{R}^n$ si A solo tiene valores propios reales) consistente de vectores propios de A.*

Demostración. Supongamos que A es diagonalizable, entonces existe una matriz invertible C tal que $C^{-1}AC = D$ o de manera equivalente $AC = CD$ con D una matriz diagonal. Sean $c_1, c_2, \ldots, c_n$ las columnas de C y sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ las entradas de la diagonal de D. Entonces

$$A\begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$\begin{bmatrix} Ac_1 & Ac_2 & \cdots & Ac_n \end{bmatrix} = \begin{bmatrix} \lambda_1 c_1 & \lambda_2 c_2 & \cdots & \lambda_n c_n \end{bmatrix}$$

Al igualar las columnas se tiene,

$$Ac_1 = \lambda_1 c_1, \quad Ac_2 = \lambda_2 c_2, \ldots, \quad Ac_n = \lambda_n c_n,$$

lo que demuestra que los vectores columna de C son vectores propios de A cuyos valores propios correspondientes son las entradas diagonales de D. Como C es invertible, vemos que las columnas de C son los n vectores propios de A linealmente independientes. Por otro lado, supongamos que $v_1, v_2, \ldots, v_n$ son n vectores propios linealmente independientes de A tales que

$$Av_i = \lambda_i v_i \quad i = 1, 2, \ldots, n. \tag{3.8}$$

Sea C la matriz cuyas columnas son $v_1, v_2, \ldots, v_n$, respectivamente. Entonces C es una matriz invertible ya que las columnas de C son linealmente independientes. Además, la ecuación (3.8) la podemos escribir como una ecuación matricial $AC = CD$, donde D es la matriz diagonal cuyas entradas de la diagonal son los valores propios de A. Como C es invertible, tenemos que $C^{-1}AC = D$, luego, A es semejante a la matriz diagonal D, es decir, A es diagonalizable. □

Teorema 3.7. *Sea A una matriz de tamaño $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ los distintos valores propios de A. Si $\mathcal{B}_i$ es una base para el espacio propio V_{λ_i}, entonces $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$, es decir, la colección total de vectores base para todos los espacios propios es linealmente independiente.*

Demostración. Sea $\mathcal{B}_i = \{v_{i1}, v_{i2}, \ldots, v_{in_i}\}$ para $i = 1, \ldots, k$. Se tiene que demostrar que $\mathcal{B} = \{v_{11}, v_{12}, \ldots, v_{1n}, v_{21}, v_{22}, \ldots, v_{2n_2}, \ldots, v_{k1}, v_{k2}, \ldots, v_{kn_k}\}$ es linealmente independiente. Supongamos que alguna combinación lineal no trivial de estos vectores es cero, por decir,

$$(c_{11}v_{11} + \cdots + c_{1n_1}v_{1n_1}) + (c_{21}v_{21} + \cdots + c_{2n_2}v_{2n_2}) + \cdots + (c_{k1}v_{k1} + \cdots + c_{kn_k}v_{kn_k}) = 0 \tag{3.9}$$

Al denotar las sumas entre paréntesis mediante $x_1, x_2, \ldots, x_k$, se puede escribir la ecuación (3.9) como

$$x_1 + x_2 + \cdots + x_k = 0 \tag{3.10}$$

Ahora cada x_i está en el espacio propio V_{λ_i} y por lo tanto es un vector propio correspondiente a λ_i ó es 0. Pero, puesto que los valores propios de λ_i son distintos, si todos los x_i son vectores propios, son linealmente independientes. Sin embargo, la ecuación (3.10) es una relación de dependencia lineal; esto es una contradicción. Se concluye que la ecuación (3.9) debe ser trivial; esto es, todo sus coeficientes son cero. Por tanto, $\mathcal{B}$ es linealmente independiente. □

Teorema 3.8. *Si A es una matriz de tamaño $n \times n$ con n valores propios distintos, entonces A es diagonalizable.*

Demostración. Sean $v_1, v_2, \ldots, v_n$ vectores propios correspondientes a los n valores propios distinos de A. Por el Teorema 3.5, $v_1, v_2, \ldots, v_n$ son linealmente independientes, de modo que, por el Teorema 3.6, A es diagonalizable. □

3.1.3 Ortogonalización

Los espacios vectoriales con los cuales estamos trabajando, son espacios vectoriales con producto interno y norma. En sí, el producto interno usual definido en $\mathbb{R}^n$ y la norma inducida por el producto interno. En este apartado mencionaremos algunos resultados sobre la ortogonalización.

Definición 3.10. *Un conjunto de vectores* $\{v_1, v_2, \ldots, v_k\}$ *en* $\mathbb{R}^n$ *se llama conjunto ortogonal si cumple que*

$$< v_i, v_j >= 0 \quad \textit{siempre que} \quad i \neq j \quad \textit{para} \quad i, j = 1, 2, \ldots, k. \tag{3.11}$$

Y es llamado conjunto ortonormal, si los vectores son unitarios, es decir, $\|v_i\|^2 = 1$ *para cada* v_i.

Teorema 3.9. *Las columnas de una matriz* Q *de tamaño* $m \times n$ *forman un conjunto ortonormal si y sólo si* $Q^T Q = I_n$.

Demostración. Es necesario demostrar que

$$(Q^T Q)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Sea q_i que denota la i-ésima columna de Q (y, en consecuencia, el i-ésimo renglón de Q^T). Dado que la entrada (i, j) de $Q^T Q$ es el producto punto del i-ésimo renglón de Q^T y la j-ésima columna de Q, se tiene que

$$(Q^T Q)_{ij} = q_i \cdot q_j \tag{3.12}$$

por la definición de multiplicación de matrices.

Ahora las columnas Q forman un conjunto ortonormal si y sólo si

$$q_i \cdot q_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

lo que, por la ecuación (3.12), se cumple si y sólo si

$$(Q^T Q)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Con esto, tenemos que las componentes de la diagonal de la matriz resultante son iguales a 1, y las componentes que no están en la diagonal son 0, por lo tanto, obtenemos la matriz identidad. □

Definición 3.11. *Una matriz* Q *de tamaño* $n \times n$ *cuyas columnas forman un conjunto ortonormal se llama matriz ortogonal.*

Teorema 3.10. *Una matriz* Q *de tamaño* $n \times n$ *es ortogonal si y sólo si* $Q^{-1} = Q^T$.

Demostración. Q es ortogonal si y sólo si $Q^T Q = I_n$, por el Teorema 3.9. Ahora consideremos la ecuación $Qx = 0$. Al multiplicar Q^T por la izquierda, se tiene $Q^T Qx = Q0 = 0$. Eso implica que $x = I_n x = 0$. Por lo tanto, el sistema representado por $Qx = 0$ tiene la solución única $x = 0$, luego, por el teorema fundamental de las matrices invertibles, consultar [Poo11] pág. 178, se sabe que Q es invertible, esto es, existe Q^{-1} y satisface $QQ^{-1} = I_n = Q^{-1}Q$. Entonces, $Q^T Q = I_n$ es verdadero si y sólo si

$$Q^T QQ^{-1} = I_n Q^{-1} \Leftrightarrow Q^T I_n = Q^{-1} \Leftrightarrow Q^T = Q^{-1}.$$

□

El siguiente teorema indica que todo subespacio de $\mathbb{R}^n$ tiene una base ortogonal; y proporciona un algoritmo para construir dicha base, tal teorema es conocido como el proceso de Gram-Schmidt.

Teorema 3.11. *Sea* $\{x_1, \ldots, x_k\}$ *una base para un subespacio* W *de* $\mathbb{R}^n$ *y defínase lo siguiente:*

$$
\begin{aligned}
v_1 &= x_1 && W_1 = \operatorname{gen}(x_1)\\
v_2 &= x_2 - \left(\frac{v_1 \cdot x_2}{v.v_1}\right)v_1 && W_2 = \operatorname{gen}(x_1, x_2)\\
v_3 &= x_3 - \left(\frac{v_1 \cdot x_3}{v.v_1}\right)v_1 - \left(\frac{v_2 \cdot x_3}{v_2 \cdot v_2}\right)v_2 && W_3 = \operatorname{gen}(x_1, x_2, x_3)\\
&\vdots\\
v_k &= x_k - \sum_{i=1}^{k-1}\left(\frac{v_i \cdot x_k}{v_i \cdot v_i}\right)v_i && W_k = \operatorname{gen}(x_1, \ldots, x_k)
\end{aligned}
$$

Entonces, para cada $i = 1, \ldots, k$, $\{v_1, \ldots, v_i\}$ *es una base ortogonal para* W_i. *En particular,* $\{v_1, \ldots v_k\}$ *es una base ortogonal para* W.

Observación 3.2. *A partir del Teorema 3.11 nótese que el primer vector de la base del subespacio* W, *es también el primer vector de la base ortogonal para* W.

Definición 3.12. *Una matriz* A *de tamaño* $n \times n$ *es diagonalizable ortogonalmente si existe una matriz ortogonal* Q *y una matriz diagonal* D *tal que* $Q^T A Q = D$.

Definición 3.13. *Si* V *es un espacio vectorial, una proyección de* V *es un operador lineal* P *sobre* V *tal que* $P^2 = P$.

3.2 TEOREMA DE DESCOMPOSICIÓN ESPECTRAL

Esta versión del Teorema espectral está dado para matrices [Poo11], uno puede hallar su equivalente para operadores lineales en algún otro texto de álgebra lineal [HK71]. A continuación enunciamos el Teorema espectral.

3.2.1 Teorema espectral

Teorema 3.12. *Sea* A *una matriz real de tamaño* $n \times n$. *Entonces* A *es simétrica si y sólo si es diagonalizable ortogonalmente.*

Demostración. Supongamos que A es diagonalizable ortogonalmente, entonces existe una matriz ortogonal Q y una matriz diagonal D tales que $Q^T A Q = D$. Dado que $Q^{-1} = Q^T$ por el Teorema 3.10, se tiene que $Q^T Q = I_n = Q Q^T$, de modo que

$$QDQ^T = Q(Q^T A Q)Q^T = (QQ^T)A(QQ^T) = I_n A I_n = A.$$

Pero entonces

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$$

pues toda matriz diagonal es simétrica. En consecuencia, A es simétrica.

Ahora, supóngase que A es simétrica. Para probar la implicación, se procederá por inducción sobre n. Para $n = 1$, no hay nada que hacer, pues una matriz de tamaño 1×1 ya está en forma diagonal. Ahora, supongamos que toda matriz simétrica con entradas reales de tamaño $k \times k$ con valores propios reales es ortogonalmente diagonalizable. Sea $n = k + 1$ y sea A una matriz simétrica de entradas reales de tamaño $n \times n$ con valores propios reales.

Sea λ_1 uno de los valores propios de A y sea v_1 un vector propio correspondiente. Entonces v_1 es un vector de entradas reales por el Teorema 3.1, y supongamos que v_1 es un vector unitario, pues de otro modo puede normalizarse y seguirá siendo un vector propio correspondiente a λ_1. Al usar el proceso de Gram-Schmidt, Teorema 3.11, podemos obtener a una base ortonormal $\{u_1, u_2, \ldots, u_n\}$ de $\mathbb{R}^n$. Ahora, se forma la matriz

$$Q_1 = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}.$$

Entonces Q_1 es ortogonal y

$$\begin{aligned} Q_1^T A Q_1 &= \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} A \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \\ &= \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} \begin{bmatrix} Au_1 & Au_2 & \cdots & Au_n \end{bmatrix} \\ &= \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} \begin{bmatrix} \lambda_1 u_1 & Au_2 & \cdots & Au_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & * \\ 0_{k\times 1} & A_1 \end{bmatrix} \\ &= B, \end{aligned}$$

dado que $u_1^T(\lambda_1 u_1) = \lambda_1(u_1^T u_1) = \lambda_1(u_1 \cdot u_1) = \lambda_1$ y $u_i^T(\lambda_1 u_1) = \lambda_1(u_i^T u_1) = \lambda_1(u_i^T u_1) = \lambda_1(u_i \cdot u_1) = 0$ para $i \neq 1$, porque $\{u_1, u_2, \ldots, u_n\}$ es un conjunto ortonormal.

Pero

$$B^T = (Q_1^T A Q_1)^T = Q_1^T A^T (Q_1^T)^T = Q_1^T A Q_1 = B$$

de modo que B es simétrica y

$$B = Q_1^{-1} A Q_1, \tag{3.13}$$

ya que Q_1 es ortogonal, ver Teorema 3.10. En consecuencia, B tiene la forma de bloque

$$B = \begin{bmatrix} \lambda_1 & 0_{1\times k} \\ 0_{k\times 1} & A_1 \end{bmatrix}$$

y A_1 es simétrica. Más aún, B es semejante a A pues se cumple la igualdad (3.13), ver Definición 3.8, de modo que el polinomio característico de B es igual al polinomio característico de A, además, el polinomio característico de A_1 divide al polinomio característico de A. Se tiene que los valores propios de A_1 tambíen son valores propios de A y, en consecuencia, son números reales; además, A_1 tiene entradas reales. Por tanto, A_1 es una matriz simétrica con entradas reales de tamaño $k \times k$ con valores propios reales, de modo que puede aplicar la hipótesis de inducción. Por lo tanto, existe una matriz ortogonal P_2 tal que $P_2^T A_1 P_2$ es una matriz diagonal, por decir D_1. Ahora sea

$$Q_2 = \begin{bmatrix} 1 & 0_{1\times k} \\ 0_{k\times 1} & P_2 \end{bmatrix}.$$

Entonces, Q_2 es una matriz ortogonal de tamaño $(k+1) \times (k+1)$, y por lo tanto, también lo es $Q = Q_1 Q_2$. Luego,

$$\begin{aligned} Q^T A Q &= (Q_1 Q_2)^T A (Q_1 Q_2) \\ &= Q_2^T (Q_1^T A Q_1) Q_2 \\ &= Q_2^T B Q_2 \\ &= \begin{bmatrix} 1 & 0 \\ 0 & P_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & P_2^T A_1 P_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0_{1\times k} \\ 0_{k\times 1} & D_1 \end{bmatrix} \\ &= D \end{aligned}$$

es decir, $Q^T AQ = D$ donde D es una matriz diagonal. Esto completa el paso de inducción y se concluye que, para todo $n \geqslant 1$ una matriz simétrica de entradas reales de tamaño $n \times n$ con valores propios reales es diagonalizable ortogonalmente. □

3.2.2 Consecuencias del teorema espectral

Observación 3.3. *Del Teorema espectral 3.12 se desprenden varias propiedades.*

i) *El Teorema espectral permite escribir una matriz simétrica de componentes reales A en la forma*

$$A = QDQ^T, \tag{3.14}$$

donde Q es una matriz ortonormal y D es la matriz diagonal. Las entradas diagonales de D son justo los valores propios de A, y las columnas de Q son los vectores ortonormales.

ii) *De i), se desprende que la matriz A puede reescribirse en su descomposición espectral*

$$A = \mu_1 P_1 + \cdots + \mu_m P_m, \tag{3.15}$$

donde P_i es la matriz de proyección ortogonal de $\mathbb{R}^n$ sobre el subespacio generado por los vectores propios ortonormales correspondientes a μ_i.

iii) *Además, la matriz de proyección P_i satisface las siguientes propiedades para $i = 1, \ldots, m$,*

- $P_i = QD_iQ^T$ *donde D_i es una matriz diagonal de tamaño $n \times n$ en cuya diagonal tiene un bloque con la matriz identidad de tamaño $m_i \times m_i$, con m_i la multiplicidad del i-ésimo valor propio μ_i,*
- $P_i^2 = P_i = P_i^T$,
- $P_iP_j = 0$ $(i \neq j)$ *donde 0 es la matriz nula de tamaño $n \times n$,*
- $I_n = P_1 + \cdots + P_m$ *donde I_n es la matriz identidad de tamaño $n \times n$,*
- $P_i = u_1^{(i)}(u_1^{(i)})^T + \ldots + u_l^{(i)}(u_l^{(i)})^T$ *donde el conjunto $\{u_1^{(i)}, \ldots, u_l^{(i)}\}$ forma una base ortonormal del i-ésimo valor propio μ_i.*

Veamos la demostración de estas propiedades.

Demostración.

i) Como A es simétrica entonces es diagonalizable ortogonalmente. Luego, existe una matriz ortonormal Q tal que $Q^TAQ = D$. Aquí las columnas de Q son vectores propios $u_1, u_2, \ldots, u_n$ ortonormales,

$$Q = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix}.$$

Luego, los vectores $u_1, u_2, \ldots, u_n$ forman un base ortonormal de los espacios propios de A. Además, Q es invertible y $Q^{-1} = Q^T$. Así,

$$\begin{aligned} Q^TAQ = D &\Rightarrow QQ^TAQQ^T = QDQ^T \\ &\Rightarrow I_nAI_n = QDQ^T \\ &\Rightarrow A = QDQ^T. \end{aligned} \tag{3.16}$$

ii) Del resultado anterior (3.16), tenemos que $A = QDQ^T$, así,

$$A = QDQ^T = \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix}.$$

En la matriz D observamos los n valores propios de la matriz A. Tales valores propios podrían repetirse, por lo que, renombramos a los valores propios como $\mu_1, \ldots, \mu_m$ como los distintos valores propios de A y con m_i, $1 \leqslant i \leqslant m$, sus respectivas multiplicidades. Ahora, podemos reescribir a D como $D = \mu_1 D_1 + \cdots + \mu_m D_m$, donde cada D_i es una matriz diagonal de tamaño $n \times n$ en cuya diagonal tiene un bloque con la matriz identidad de tamaño $m_i \times m_i$, esto es,

$$D_i = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & I_{m_i} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}.$$

Luego,

$$\begin{aligned} A &= QDQ^T \\ &= Q(\mu_1 D_1 + \cdots + \mu_m D_m)Q^T \\ &= \mu_1 QD_1Q^T + \cdots + \mu_m QD_mQ^T \\ &= \mu_1 P_1 + \cdots + \mu_m P_m. \end{aligned}$$

Entonces A tiene la descomposición espectral

$$A = \mu_1 P_1 + \ldots + \mu_m P_m,$$

donde $P_i = QD_iQ^T$ $i = 1, \ldots, m$, y esta matriz es llamada la matriz de proyección ortogonal sobre el subespacio vectorial generado por los m_i vectores propios ortonormales correspondientes al valor propio μ_i.

iii) Del inciso anterior observamos que se satisface que $P_i = QD_iQ^T$, para cada $i = 1, \ldots, m$.

Ahora, veamos que se cumple que para cada $i = 1, \ldots, m$, $P_i^2 = P_i$, $P_i^T = P_i$, $P_iP_j = 0$ con $i \neq j$, $P_1 + \cdots + P_k = I_n$ donde I_n es la matriz identidad de tamaño $n \times n$, y además que, $P_i = u_1^{(i)}(u_1^{(i)})^T + \cdots + u_l^{(i)}(u_l^{(i)})^T$.

Como,

$$\begin{aligned} P_i^2 &= P_iP_i \\ &= (QD_iQ^T)(QD_iQ^T) \\ &= QD_i(Q^TQ)D_iQ^T \\ &= QD_iI_nD_iQ^T \\ &= Q(D_iD_i)Q^T \\ &= QD_i^2Q^T \\ &= QD_iQ^T \\ &= P_i, \end{aligned}$$

pues recordar que $Q^TQ = I_n$ donde I_n es la matriz identidad de tamaño $n \times n$ por el Teorema 3.9, y D_i es una matriz diagonal que tiene un bloque con la matriz identidad y cualquier potencia de la matriz identidad sigue siendo la matriz identidad, por lo que $D_i^2 = D_i$. Luego, se cumple que $P_i^2 = P_i$.

Además,

$$\begin{aligned} P_i^T &= (QD_iQ^T)^T \\ &= (Q^T)^T(D_i)^T(Q)^T \\ &= QD_iQ^T \\ &= P_i, \end{aligned}$$

ya que D_i es una matriz en cuya diagonal tiene un bloque con la matriz identidad y el resto de sus componentes es 0 se cumple que es una matriz simétrica, y así $D_i^T = D_i$. Por lo tanto, $P_i^T = P_i$.

Ahora, sea $i \neq j$

$$\begin{aligned} P_iP_j &= (QD_iQ^T)(QD_jQ^T) \\ &= QD_i(Q^TQ)D_jQ^T \\ &= QD_iI_nD_jQ^T \\ &= Q(D_iD_j)Q^T \\ &= Q0Q^T \\ &= 0, \end{aligned}$$

pues $Q^TQ = I_n$ por el Teorema 3.9. D_i, D_j son matrices en cuya diagonal tienen un bloque de la matriz identidad de tamaño $m_i \times m_i$ y $m_j \times m_j$, respectivamente; que además ocupan diferentes entradas y el resto de sus entradas son 0. Por lo que, al realizar el producto obtenemos la matriz nula. Y así, $P_iP_j = 0$ con $i \neq j$.

Además,

$$\begin{aligned} P_1 + P_2 + \cdots + P_m &= QD_1Q^T + QD_2Q^T + \cdots + QD_mQ^T \\ &= Q(D_1 + D_2 + \cdots + D_m)Q^T \\ &= QI_nQ^T \\ &= QQ^T \\ &= I_n, \end{aligned}$$

pues cada D_i es una matriz en cuya diagonal tiene un bloque con la matriz identidad de tamaño $m_i \times m_i$ que son las multiplicidades de cada valor propio, además, ningún bloque coincide con la posición del otro. Entonces al sumarlos obtenemos la matriz diagonal cuyas componentes en la diagonal es 1 y el resto de las componentes es 0, es decir, tenemos una matriz identidad de tamaño $n \times n$. Además, por el Teorema 3.10 $Q^T = Q^{-1}$, entonces $QQ^T = QQ^{-1} = I_n$ donde I_n es la matriz identidad de tamaño $n \times n$. Por lo tanto, $P_1 + P_2 + \cdots + P_m = I_n$.

Ahora, para cada valor propio tenemos una base ortonormal para el subespacio vectorial correspondiente, y supongamos que el subespacio vectorial V_{μ_i} correspondiente al i-ésimo valor propio μ_i tiene como base ortonormal al conjunto $\{u_1^{(i)}, \ldots, u_l^{(i)}\}$, entonces

$$\begin{aligned} P_i &= QD_iQ^T \\ &= \begin{bmatrix} u_1^{(1)} & \cdots & u_k^{(1)} & \cdots & u_1^{(i)} & \cdots & u_l^{(i)} & \cdots & u_1^{(m)} & \cdots & u_n^{(m)} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{m_i} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (u_1^{(1)})^T \\ \vdots \\ (u_k^{(1)})^T \\ \vdots \\ (u_1^{(i)})^T \\ \vdots \\ (u_l^{(i)})^T \\ \vdots \\ (u_1^{(m)})^T \\ \vdots \\ (u_n^{(m)})^T \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{bmatrix} 0 \cdots 0 \cdots u_{11}^{(i)} \cdots u_{l1}^{(i)} \cdots 0 \cdots 0 \\ 0 \cdots 0 \cdots u_{12}^{(i)} \cdots u_{l2}^{(i)} \cdots 0 \cdots 0 \\ \vdots \ddots \vdots \ddots \vdots \ddots \vdots \ddots \vdots \ddots \vdots \\ 0 \cdots 0 \cdots u_{1n}^{(i)} \cdots u_{ln}^{(i)} \cdots 0 \cdots 0 \end{bmatrix} \begin{bmatrix} u_{11}^{(1)} & u_{12}^{(1)} & \cdots & u_{1n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k1}^{(1)} & u_{k2}^{(1)} & \cdots & u_{kn}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{11}^{(i)} & u_{12}^{(i)} & \cdots & u_{1n}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{l1}^{(i)} & u_{l2}^{(i)} & \cdots & u_{ln}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{11}^{(m)} & u_{12}^{(m)} & \cdots & u_{1n}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1}^{(m)} & u_{n2}^{(m)} & \cdots & u_{nn}^{(m)} \end{bmatrix} \\
&= \begin{bmatrix} (u_{11}^{(i)})^2 + \cdots + (u_{l1}^{(i)})^2 & u_{11}^{(i)}u_{12}^{(i)} + \cdots + u_{l1}^{(i)}u_{l2}^{(i)} & \cdots & u_{11}^{(i)}u_{1n}^{(i)} + \cdots + u_{l1}^{(i)}u_{ln}^{(i)} \\ u_{12}^{(i)}u_{11}^{(i)} + \cdots + u_{l2}^{(i)}u_{l1}^{(i)} & (u_{12}^{(i)})^2 + \cdots + (u_{l2}^{(i)})^2 & \cdots & u_{12}^{(i)}u_{1n}^{(i)} + \cdots + u_{l2}^{(i)}u_{ln}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n}^{(i)}u_{11}^{(i)} + \cdots + u_{ln}^{(i)}u_{l1}^{(i)} & u_{1n}^{(i)}u_{12}^{(i)} + \cdots + u_{ln}^{(i)}u_{l2}^{(i)} & \cdots & (u_{1n}^{(i)})^2 + \cdots + (u_{ln}^{(i)})^2 \end{bmatrix} \\
&= \begin{bmatrix} (u_{11}^{(i)})^2 & u_{11}^{(i)}u_{12}^{(i)} & \cdots & u_{11}^{(i)}u_{1n}^{(i)} \\ u_{12}^{(i)}u_{11}^{(i)} & (u_{12}^{(i)})^2 & \cdots & u_{12}^{(i)}u_{1n}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n}^{(i)}u_{11}^{(i)} & u_{1n}^{(i)}u_{12}^{(i)} & \cdots & (u_{1n}^{(i)})^2 \end{bmatrix} + \cdots + \begin{bmatrix} (u_{l1}^{(i)})^2 & u_{l1}^{(i)}u_{l2}^{(i)} & \cdots & u_{l1}^{(i)}u_{ln}^{(i)} \\ u_{l2}^{(i)}u_{l1}^{(i)} & (u_{l2}^{(i)})^2 & \cdots & u_{l2}^{(i)}u_{ln}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{ln}^{(i)}u_{l1}^{(i)} & u_{ln}^{(i)}u_{l2}^{(i)} & \cdots & (u_{ln}^{(i)})^2 \end{bmatrix} \\
&= \begin{bmatrix} u_{11}^{(i)} \\ u_{12}^{(i)} \\ \vdots \\ u_{1n}^{(i)} \end{bmatrix} \begin{bmatrix} u_{11}^{(i)} & u_{12}^{(i)} & \cdots & u_{1n}^{(i)} \end{bmatrix} + \cdots + \begin{bmatrix} u_{l1}^{(i)} \\ u_{l2}^{(i)} \\ \vdots \\ u_{ln}^{(i)} \end{bmatrix} \begin{bmatrix} u_{l1}^{(i)} & u_{l2}^{(i)} & \cdots & u_{ln}^{(i)} \end{bmatrix} \\
&= u_1(u_1^{(i)})^T + \cdots + u_l(u_l^{(i)})^T,
\end{aligned}
$$

donde u_i es el vector ortonormal columna de tamaño $n \times 1$ y al realizar el producto con su transpuesto obtenemos una matriz de tamaño $n \times n$. Por lo tanto, para cada $i = 1, \ldots, m$ $P_i = u_1(u_1^{(i)})^T + \cdots + u_l(u_l^{(i)})^T$.

□

3.2.3 Ejemplos

Veamos en los siguientes ejemplos cómo diagonalizar ortogonalmente una matriz y cómo se cumplen las propiedades mencionadas de la Observación 3.3.

Ejemplo 3.1. *Sea $\mathcal{C}(\mathcal{B})$ el $[8,5,2]$-código LDPC $(2,4)$-regular con matriz de chequeo de paridad H y grafo bipartito $\mathcal{B}$ del Ejemplo 4.1. Considerando su matriz H vamos a obtener la matriz H^TH que resulta ser una matriz de valores reales simétrica de tamaño 8×8, la cual de acuerdo al Teorema 3.12 es diagonalizable ortogonalmente. Para comprobar esto debemos hallar los valores propios, los vectores propios, diagonalizar, calcular los vectores ortonormales y diagonalizar ortogonalmente; y después obtendremos las matrices de proyección para escribir a la matriz H^TH en su descomposición matricial. Con ayuda del software **wxMaxima 16.04.2** [VLIK15] vamos a realizar los cálculos necesarios.*

```
(%i1)  H: matrix( [0,1,0,1,1,0,0,1], [1,1,1,0,0,1,0,0], [0,0,1,0,0,1,1,1], [1,0,0,1,1,0,1,0] )$
```

(%i2) H; (Matriz de chequeo de paridad del código.)

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix} \tag{\%o2}$$

(%i3) HT: transpose(H); (Matriz transpuesta de la matriz H.)

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

(%i4) B: HT.H; (Veremos si esta matriz de tamaño 8×8 es diagonalizable ortogonalmente.)

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

(%i5) pt: charpoly(B, t), expand; (Polinomio característico de la matriz B.)

$$t^8 - 16t^7 + 84t^6 - 176t^5 + 128t^4 \tag{pt}$$

(%i6) factor(pt); (Factorización del polinomio característico.)

$$(t-8)\,(t-4)\,(t-2)^2\,t^4 \tag{\%o6}$$

(%i7) eigenvalues(B); (Valores propios de la matriz B con sus respectivas multiplicidades.)

$$[[4, 8, 2, 0], [1, 1, 2, 4]] \tag{\%o7}$$

(Espacio propio correspondiente al valor propio 8 con multiplicidad 1.)

(%i8) C1: B - 8*ident(8);

$$\begin{pmatrix} -6 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & -6 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & -6 & 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 0 & -6 & 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & -6 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 & 0 & -6 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & -6 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & -6 \end{pmatrix} \tag{C1}$$

```
(%i9)  linsolve([
-6*x1+x2+x3+x4+x5+x6+x7=0,
x1-6*x2+x3+x4+x5+x6+x8=0,
x1+x2-6*x3+2*x6+x7+x8=0,
x1+x2-6*x4+2*x5+x7+x8=0,
x1+x2+2*x4-6*x5+x7+x8=0,
x1+x2+2*x3-6*x6+x7+x8=0,
x1+x3+x4+x5+x6-6*x7+x8=0,
x2+x3+x4+x5+x6+x7-6*x8=0],
[x1,x2,x3,x4,x5,x6,x7,x8] );
```

solve: dependent equations eliminated: (8)

$$[x1 = \%r2, x2 = \%r2, x3 = \%r2, x4 = \%r2, x5 = \%r2, x6 = \%r2, x7 = \%r2, x8 = \%r2] \qquad (\%o9)$$

(Espacio propio correspondiente al valor propio 4 con multiplicidad 1.)

```
(%i10)  C2: B - 4*ident(8);
```

$$\begin{pmatrix} -2 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & -2 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & -2 & 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 0 & -2 & 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & -2 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 & 0 & -2 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & -2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & -2 \end{pmatrix} \qquad (C2)$$

```
(%i11)  linsolve([
-2*x1+x2+x3+x4+x5+x6+x7=0,
x1-2*x2+x3+x4+x5+x6+x8=0,
x1+x2-2*x3+2*x6+x7+x8=0,
x1+x2-2*x4+2*x5+x7+x8=0,
x1+x2+2*x4-2*x5+x7+x8=0,
x1+x2+2*x3-2*x6+x7+x8=0,
x1+x3+x4+x5+x6-2*x7+x8=0,
x2+x3+x4+x5+x6+x7-2*x8=0],
[x1,x2,x3,x4,x5,x6,x7,x8] );
```

solve: dependent equations eliminated: (5)

$$[x1 = 0, x2 = 0, x3 = \%r1, x4 = -\%r1, x5 = -\%r1, x6 = \%r1, x7 = 0, x8 = 0] \qquad (\%o11)$$

(Espacio propio correspondiente al valor propio 2 con multiplicidad 2.)

```
(%i12)  C3: B - 2*ident(8);
```

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 & 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \qquad (C3)$$

```
(%i13) linsolve([
-0*x1+x2+x3+x4+x5+x6+x7=0,
x1-0*x2+x3+x4+x5+x6+x8=0,
x1+x2-0*x3+2*x6+x7+x8=0,
x1+x2-0*x4+2*x5+x7+x8=0,
x1+x2+2*x4-0*x5+x7+x8=0,
x1+x2+2*x3-0*x6+x7+x8=0,
x1+x3+x4+x5+x6-0*x7+x8=0,
x2+x3+x4+x5+x6+x7-0*x8=0],
[x1,x2,x3,x4,x5,x6,x7,x8] );
```

solve: dependent equations eliminated: (8 7)

$$[x1 = -\%r3, x2 = -\%r4, x3 = 0, x4 = 0, x5 = 0, x6 = 0, x7 = \%r4, x8 = \%r3] \qquad (\%o13)$$

(Espacio propio correspondiente al valor propio 0 con multiplicidad 4.)

```
(%i14) C4: B - 0*ident(8);
```

$$\begin{pmatrix} 2 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 2 & 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 & 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 2 \end{pmatrix} \qquad (C4)$$

```
(%i15) linsolve([
2*x1+x2+x3+x4+x5+x6+x7=0,
x1+2*x2+x3+x4+x5+x6+x8=0,
x1+x2+2*x3+2*x6+x7+x8=0,
x1+x2+2*x4+2*x5+x7+x8=0,
x1+x2+2*x4+2*x5+x7+x8=0,
x1+x2+2*x3+2*x6+x7+x8=0,
x1+x3+x4+x5+x6+2*x7+x8=0,
x2+x3+x4+x5+x6+x7+2*x8=0],
[x1,x2,x3,x4,x5,x6,x7,x8] );
```

solve: dependent equations eliminated: (6 5 7 8)

$$[x1 = \%r5, x2 = \%r6, x3 = -\%r7 - \%r6 - \%r5, x4 = -\%r8 - \%r6 - \%r5, x5 = \%r8, x6 = \%r7, x7 = \%r6, x8 = \%r5] \qquad (\%o15)$$

```
(%i16) XT: matrix( [1,1,1,1,1,1,1,1], [0,0,1,-1,-1,1,0,0], [-1,0,0,0,0,0,0,1], [0,-1,0,0,0,0,1,0],
[1,0,-1,-1,0,0,0,1], [0,1,-1,-1,0,0,1,0], [0,0,-1,0,0,1,0,0], [0,0,0,-1,1,0,0,0] )$
(%i17) XT;   (Matriz cuyos renglones son vectores propios linealmente independientes de la matriz B.)
```

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix} \qquad (\%o17)$$

(%i18) X: transpose(XT); (Matriz cuyas columnas son vectores propios linealmente independientes de la matriz B.)

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 & -1 & 0 \\ 1 & -1 & 0 & 0 & -1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \tag{X}$$

(%i19) Y: invert(X); (Matriz inversa de la matriz X.)

$$\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \end{pmatrix} \tag{Y}$$

(%i20) D: matrix([8,0,0,0,0,0,0,0], [0,4,0,0,0,0,0,0], [0,0,2,0,0,0,0,0], [0,0,0,2,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i21) D; (Matriz diagonal cuya diagonal está formada por los valores propios de la matriz B.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{D}$$

(%i22) D: Y.B.X; (Comprobamos que B es semejante a D, y como D es una matriz diagonal entonces B es diagonalizable.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o21}$$

(%i23) load (eigen)$ (Cargamos este paquete para poder ejecutar la instrucción del proceso de Gram-Schmidt.)

(%i24) Z: gramschmidt (XT); (Proceso de Gram-Schmidt que construye una base ortogonal.)

$$[[1,1,1,1,1,1,1,1],[0,0,1,-1,-1,1,0,0],[-1,0,0,0,0,0,0,1],[0,-1,0,0,0,0,1,0],[1,0,-1,-1,0,0,0,1],$$

$$[-\frac{1}{2},1,-\frac{1}{2},-\frac{1}{2},0,0,1,-\frac{1}{2}],[-\frac{1}{23},-\frac{1}{23},-\frac{2}{3},\frac{1}{3},0,1,-\frac{1}{23},-\frac{1}{23}],[-\frac{1}{5},-\frac{1}{5},\frac{1}{5},-\frac{3}{5},1,\frac{1}{5},-\frac{1}{5},-\frac{1}{5}]] \tag{Z}$$

(%i25) ratsimp(Z); (Simplificamos la salida anterior.)

$$[[1,1,1,1,1,1,1,1],[0,0,1,-1,-1,1,0,0],[-1,0,0,0,0,0,0,1],[0,-1,0,0,0,0,1,0],[1,0,-1,-1,0,0,0,1],$$

$$[-\frac{1}{2},1,-\frac{1}{2},-\frac{1}{2},0,0,1,-\frac{1}{2}],[-\frac{1}{6},-\frac{1}{6},-\frac{2}{3},\frac{1}{3},0,1,-\frac{1}{6},-\frac{1}{6}],[-\frac{1}{5},-\frac{1}{5},\frac{1}{5},-\frac{3}{5},1,\frac{1}{5},-\frac{1}{5},-\frac{1}{5}]] \quad (\%o25)$$

(%i26)
```
QT: matrix(
(1/(2*sqrt(2)))*[1,1,1,1,1,1,1,1],
(1/2)*[0,0,1,-1,-1,1,0,0],
(1/sqrt(2))*[-1,0,0,0,0,0,0,1],
(1/sqrt(2))*[0,-1,0,0,0,0,1,0],
(1/2)*[1,0,-1,-1,0,0,0,1],
(1/sqrt(3))*[-1/2,1,-1/2,-1/2,0,0,1,-1/2],
(sqrt(9)/sqrt(15))*[-1/6,-1/6,-2/3,1/3,0,1,-1/6,-1/6],
(5/sqrt(40))*[-1/5,-1/5,1/5,-3/5,1,1/5,-1/5,-1/5] )$
```

(%i27) QT; (Matriz cuyos renglones son los vectores ortogonales normalizados.)

$$\begin{pmatrix} \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{15}} & -\frac{2}{\sqrt{15}} & \frac{1}{\sqrt{15}} & 0 & \frac{3}{\sqrt{15}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{15}} \\ -\frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} & \frac{1}{2\sqrt{10}} & -\frac{3}{2\sqrt{10}} & \frac{5}{2\sqrt{10}} & \frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} \end{pmatrix} \quad (\%o27)$$

(%i28) Q: transpose(QT); (Matriz cuyas columnas son los vectores ortonormales calculados.)

$$\begin{pmatrix} \frac{1}{2^{\frac{3}{2}}} & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{2}{\sqrt{15}} & \frac{1}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{15}} & -\frac{3}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{5}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{3}{\sqrt{15}} & \frac{1}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ \frac{1}{2^{\frac{3}{2}}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \end{pmatrix} \quad (Q)$$

(%i29) D: QT.B.Q; (Comprobamos que la matriz B es diagonalizable ortogonalmente.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & \frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2^{\frac{3}{2}}}+\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2^{\frac{3}{2}}} & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & \frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{\sqrt{2}}-\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{\sqrt{3}}+\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2\sqrt{15}}-\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2\sqrt{15}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2\sqrt{10}}-\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{2\sqrt{10}} & 0 & 0 \end{pmatrix} \quad (D)$$

(%i30) ratsimp(D); (Simplificamos la salida anterior.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\%o30)$$

Hasta aquí hemos comprobado que la matriz $H^T H$ *es diagonalizable ortogonalmente. En lo que sigue comprobaremos que se satisfacen las propiedades de la Observación 3.3 que son consecuencias del Teorema espectral.*

(%i31) B: Q.D.QT; (La matriz B puede escribirse en la forma QDQ^T.)

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

(%i32) D1: matrix([1,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i33) D1; (Construimos la matriz asociada al valor propio 8 con multiplicidad 1 cuya diagonal tiene un bloque con la matriz identidad de tamaño 1×1 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\%o33)$$

(%i34) P1: Q.D1.QT; (Matriz de proyección del espacio propio de B correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \quad (P1)$$

(%i35) D2: matrix([0,0,0,0,0,0,0,0], [0,1,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i36) D2; (Construimos la matriz asociada al valor propio 4 con multiplicidad 1 cuya diagonal tiene un bloque con la matriz identidad de tamaño 1×1 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o36}$$

(%i37) P2: Q.D2.QT; (Matriz de proyección del espacio propio de B correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{P2}$$

(%i38) D3: matrix([0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,1,0,0,0,0,0], [0,0,0,1,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i39) D3; (Construimos la matriz asociada al valor propio 2 con multiplicidad 2 cuya diagonal tiene un bloque con la matriz identidad de tamaño 2×2 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o39}$$

(%i40) P3: Q.D3.QT; (Matriz de proyección del espacio propio de B correspondiente al espacio propio 2.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3}$$

(%i41) D4: matrix([0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,1,0,0,0], [0,0,0,0,0,1,0,0], [0,0,0,0,0,0,1,0], [0,0,0,0,0,0,0,1])$

(%i42) D4; (Construimos la matriz asociada al valor propio 0 con multiplicidad 4 cuya diagonal tiene un bloque con la matriz identidad de tamaño 4×4 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{%o42}$$

(%i43) P4: Q.D4.QT; (Matriz de proyección del espacio propio de B correspondiente al valor propio 0.)

$$\begin{pmatrix} \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \end{pmatrix} \tag{P4}$$

(%i44) D: 8*D1 + 4*D2 + 2*D3 + 0*D4; (Forma en la que descomponemos a la matriz D.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{D}$$

(%i45) P1.P1; (Comprobamos que la matriz de proyección P1 elevada al cuadrado es igual a P1.)

$$\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \tag{%o44}$$

(%i46) P2.P2; (Comprobamos que la matriz de proyección P2 elevada al cuadrado es igual a P2.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{%o46}$$

(%i47) P3.P3; (Comprobamos que la matriz de proyección P3 elevada al cuadrado es igual a P3.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{\%o47}$$

(%i48) P4.P4; (Comprobamos que la matriz de proyección P elevada al cuadrado es igual a P4.)

$$\begin{pmatrix} \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \end{pmatrix} \tag{\%o48}$$

(%i49) P1T: transpose(P1); (Comprobamos que la matriz de proyección P1 transpuesta es igual a P1.)

$$\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \tag{P1T}$$

(%i50) P2T: transpose(P2); (Comprobamos que la matriz de proyección P2 transpuesta es igual a P2.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{P2T}$$

(%i51) P3T: transpose(P3); (Comprobamos que la matriz de proyección P3 transpuesta es igual a P3.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3T}$$

(%i52) P4T: transpose(P4); (Comprobamos que la matriz de proyección P4 transpuesta es igual a P4.)

$$\begin{pmatrix} \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \end{pmatrix} \tag{P4T}$$

(%i53) P1.P2; (Al multiplicar las matrices de proyección P1, P2 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{%o53}$$

(%i54) P1.P3; (Al multiplicar las matrices de proyección P1, P3 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{%o54}$$

(%i55) P1.P4; (Al multiplicar las matrices de proyección P1, P4 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{%o55}$$

(%i56) P2.P1; (Al multiplicar las matrices de proyección P2, P1 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{%o56}$$

(%i57) P2.P3; (Al multiplicar las matrices de proyección P2, P3 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o57}$$

(%i58) P2.P4; (Al multiplicar las matrices de proyección P2, P4 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o58}$$

(%i59) P3.P1; (Al multiplicar las matrices de proyección P3, P1 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o59}$$

(%i60) P3.P2; (Al multiplicar las matrices de proyección P3, P2 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o60}$$

(%i61) P3.P4; (Al multiplicar las matrices de proyección P3, P4 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o61}$$

(%i62) P4.P1; (Al multiplicar las matrices de proyección P4, P1 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o62}$$

(%i63) P4.P2; (Al multiplicar las matrices de proyección P4, P2 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o63}$$

(%i64) P4.P3; (Al multiplicar las matrices de proyección P4, P3 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o64}$$

(%i65) I: P1 + P2 + P3 + P4; (Comprobamos que la suma de las matrices de proyección es igual a la matriz identidad $I_{8\times 8}$.)

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

(%i66) U1T: matrix((1/(2*sqrt(2)))*[1,1,1,1,1,1,1,1], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i67) U1T; (Matriz cuyo primer renglón es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} & \frac{1}{2^{\frac{3}{2}}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o67}$$

(%i68) U1: transpose(U1T); (Matriz cuya primer columna es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2^{\frac{3}{2}}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{U1}$$

(%i69) P1: U1.U1T; (Otra forma de obtener la matriz de proyección P1 es multiplicando las matrices formadas por el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \tag{P1}$$

(%i70) U2T: matrix([0,0,0,0,0,0,0,0], (1/2)*[0,0,1,-1,-1,1,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i71) U2T; (Matriz cuyo segundo renglón es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o71}$$

(%i72) U2: transpose(U2T); (Matriz cuya segunda columna es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{U2}$$

(%i73) P2: U2.U2T; (Otra forma de obtener la matriz de proyección P2 es multiplicando las matrices formadas por el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{P2}$$

(%i74) U3T: matrix([0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], (1/sqrt(2))*[-1,0,0,0,0,0,0,1], (1/sqrt(2))*[0,-1,0,0,0,0,1,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0])$

(%i75) U3T; (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o75}$$

(%i76) U3: transpose(U3T); (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{U3}$$

(%i77) P3: U3.U3T; (Otra forma de obtener la matriz de proyección P3 es multiplicando las matrices formadas por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3}$$

(%i78) U4T: matrix([0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0], (1/2)*[1,0,-1,-1,0,0,0,1], (1/sqrt(3))*[-1/2,1,-1/2,-1/2,0,0,1,-1/2], (sqrt(9)/sqrt(15))*[-1/6,-1/6,-2/3,1/3,0,1,-1/6,-1/6], (5/sqrt(40))*[-1/5,-1/5,1/5,-3/5,1,1/5,-1/5,-1/5])$

(%i79) U4T; (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 0.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{15}} & -\frac{2}{\sqrt{15}} & \frac{1}{\sqrt{15}} & 0 & \frac{3}{\sqrt{15}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{15}} \\ -\frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} & \frac{1}{2\sqrt{10}} & -\frac{3}{2\sqrt{10}} & \frac{5}{2\sqrt{10}} & \frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} & -\frac{1}{2\sqrt{10}} \end{pmatrix} \tag{\%o79}$$

(%i80) U4: transpose(U4T); (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 0.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{2}{\sqrt{15}} & \frac{1}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{15}} & -\frac{3}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{\sqrt{15}} & \frac{1}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{15}} & -\frac{1}{2\sqrt{10}} \end{pmatrix} \tag{U4}$$

(%i81) P4: U4.U4T; (Otra forma de obtener la matriz de proyección P4 es multiplicando las matrices formadas por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 0.)

$$\begin{pmatrix} \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{5}{8} & \frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{3}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} & -\frac{3}{8} & \frac{5}{8} & \frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & -\frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} & -\frac{1}{8} \\ \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{8} & \frac{3}{8} \end{pmatrix} \tag{P4}$$

(%i82) B: 8*P1 + 4*P2 + 2*P3 + 0*P4; (Obtenemos la descomposición espectral de la matriz B.)

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

Acabamos de mostrar que se cumplen todas las propiedades del Teorema de descomposición espectral para la matriz $H^T H$.

Ejemplo 3.2. *Sea $\mathcal{C}(\mathcal{B})$ el $[8,5,2]$-código LDPC $(2,4)$-regular con matriz de chequeo de paridad H y grafo bipartito $\mathcal{B}$ del Ejemplo 4.1. Considerando su matriz H veamos que ocurre con la matriz HH^T que resulta ser una matriz de valores reales simétrica de tamaño 4×4, la cual de acuerdo al Teorema 3.12 es diagonalizable ortogonalmente. Para comprobar esto debemos hallar los valores propios, los vectores propios, diagonalizar, calcular los vectores ortonormales y diagonalizar ortogonalmente; y después obtendremos las matrices de proyección para escribir a la matriz HH^T en su descomposición matricial. Nuevamente usando del software* ***wxMaxima 16.04.2*** *[VLIK15] vamos a realizar los cálculos necesarios.*

```
(%i1)  H: matrix( [0,1,0,1,1,0,0,1], [1,1,1,0,0,1,0,0], [0,0,1,0,0,1,1,1], [1,0,0,1,1,0,1,0] )$
(%i2)  H;   (Matriz de chequeo de paridad del código.)
```

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix} \tag{\%o2}$$

```
(%i3)  HT: transpose(H);   (Matriz transpuesta de la matriz H.)
```

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

```
(%i4)  C: H.HT;   (Veremos si esta matriz de tamaño 4 × 4 es diagonalizable ortogonalmente.)
```

$$\begin{pmatrix} 4 & 1 & 1 & 2 \\ 1 & 4 & 2 & 1 \\ 1 & 2 & 4 & 1 \\ 2 & 1 & 1 & 4 \end{pmatrix} \tag{C}$$

```
(%i5)  pt: charpoly(C, t), expand;   (Polinomio característico de la matriz C.)
```

$$t^4 - 16t^3 + 84t^2 - 176t + 128 \tag{pt}$$

```
(%i6)  factor(pt);   (Factorización del polinomio característico.)
```

$$(t-8)\,(t-4)\,(t-2)^2 \tag{\%o6}$$

```
(%i7)  eigenvalues(C);   (Valores propios de la matriz C con sus respectivas multiplicidades.)
```

$$[[4,8,2],[1,1,2]] \tag{\%o7}$$

(Espacio propio correspondiente al valor propio 8 con multiplicidad 1.)

```
(%i8)  C1: C - 8*ident(4);
```

$$\begin{pmatrix} -4 & 1 & 1 & 2 \\ 1 & -4 & 2 & 1 \\ 1 & 2 & -4 & 1 \\ 2 & 1 & 1 & -4 \end{pmatrix} \tag{C1}$$

```
(%i9)  linsolve([
       -4*x1+x2+x3+2*x4=0,
       x1-4*x2+2*x3+x4=0,
       x1+2*x2-4*x3+x4=0,
       2*x1+x2+x3-4*x4=0], [x1,x2,x3,x4]);
```

solve: dependent equations eliminated: (4)

$$[x1 = \%r1, x2 = \%r1, x3 = \%r1, x4 = \%r1] \tag{\%o9}$$

(Espacio propio correspondiente al valor propio 4 con multiplicidad 1.)

(%i10) C2: C - 4*ident(4);

$$\begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \tag{C2}$$

(%i11) linsolve([x2+x3+2*x4=0, x1+2*x3+x4=0, x1+2*x2+x4=0, 2*x1+x2+x3=0], [x1,x2,x3,x4]);

solve: dependent equations eliminated: (2)

$$[x1 = \%r2, x2 = -\%r2, x3 = -\%r2, x4 = \%r2] \tag{\%o11}$$

(Espacio propio correspondiente al valor propio 2 con multiplicidad 2.)

(%i12) C3: C - 2*ident(4);

$$\begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix} \tag{C3}$$

(%i13) linsolve([
2*x1+x2+x3+2*x4=0,
x1+2*x2+2*x3+x4=0,
x1+2*x2+2*x3+x4=0,
2*x1+x2+x3+2*x4=0], [x1,x2,x3,x4]);

solve: dependent equations eliminated: (4 3)

$$[x1 = -\%r3, x2 = -\%r4, x3 = \%r4, x4 = \%r3] \tag{\%o13}$$

(%i14) XT: matrix([1,1,1,1], [1,-1,-1,1], [-1,0,0,1], [0,-1,1,0])$

(%i15) XT; (Matriz cuyos renglones son vectores propios linealmente independientes de la matriz C.)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix} \tag{\%o15}$$

(%i16) X: transpose(XT); (Matriz cuyas columnas son vectores propios linealmente independientes de la matriz C.)

$$\begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \tag{X}$$

(%i17) Y: invert(X); (Matriz inversa de la matriz X.)

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \tag{Y}$$

(%i18) D: matrix([8,0,0,0], [0,4,0,0], [0,0,2,0], [0,0,0,2])$

(%i19) D; (Matriz diagonal cuya diagonal está formada por los valores propios de la matriz C.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \tag{\%o19}$$

(%i20) D: Y.C.X; (Comprobamos que C es semejante a D, y como D es una matriz diagonal entonces C es diagonalizable.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \tag{\%o19}$$

(%i21) load (eigen)$ (Cargamos este paquete para poder ejecutar la instrucción del proceso de Gram-Schmidt.)

(%i22) Z: gramschmidt (XT); (Proceso de Gram-Schmidt que construye una base ortogonal.)

$$[[1,1,1,1],[1,-1,-1,1],[-1,0,0,1],[0,-1,1,0]] \tag{Z}$$

(%i23) QT: matrix((1/2)*[1,1,1,1], (1/2)*[1,-1,-1,1], (1/sqrt(2))*[-1,0,0,1], (1/sqrt(2))*[0,-1,1,0])$

(%i24) QT; (Matriz cuyos renglones son los vectores ortogonales normalizados.)

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \tag{QT}$$

(%i25) Q: transpose(QT); (Matriz cuyas columnas son los vectores ortonormales.)

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \tag{Q}$$

(%i26) D: QT.C.Q; (Comprobamos que la matriz C es diagonalizable ortogonalmente.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \tag{D}$$

Hasta aquí hemos comprobado que la matriz HH^T *es diagonalizable ortogonalmente. En lo que sigue comprobaremos que se satisfacen las propiedades de la Observación 3.3 que son consecuencias del Teorema espectral.*

(%i27) C: Q.D.QT; (La matriz C puede escribirse en la forma QDQ^T.)

$$\begin{pmatrix} 4 & 1 & 1 & 2 \\ 1 & 4 & 2 & 1 \\ 1 & 2 & 4 & 1 \\ 2 & 1 & 1 & 4 \end{pmatrix} \tag{C}$$

(%i28) D1: matrix([1,0,0,0], [0,0,0,0], [0,0,0,0], [0,0,0,0])$

(%i29) D1; (Construimos la matriz asociada al valor propio 8 con multiplicidad 1 cuya diagonal tiene un bloque con la matriz identidad de tamaño 1×1 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o29}$$

(%i30) P1: Q.D1.QT; (Matriz de proyección del espacio propio de C correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P1}$$

(%i31) D2: matrix([0,0,0,0], [0,1,0,0], [0,0,0,0], [0,0,0,0])$

(%i32) D2; (Construimos la matriz asociada al valor propio 4 con multiplicidad 1 cuya diagonal tiene un bloque con la matriz identidad de tamaño 1×1 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o32}$$

(%i33) P2: Q.D2.QT; (Matriz de proyección del espacio propio de C correspondiente al valor propio 4.)

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P2}$$

(%i34) D3: matrix([0,0,0,0], [0,0,0,0], [0,0,1,0], [0,0,0,1])$

(%i35) D3; (Construimos la matriz asociada al valor propio 2 con multiplicidad 2 cuya diagonal tiene un bloque con la matriz identidad de tamaño 2×2 y el resto de sus elementos son ceros.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{\%o35}$$

(%i36) P3: Q.D3.QT; (Matriz de proyección del espacio propio de C correspondiente al valor propio 2.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3}$$

(%i37) D: 8*D1 + 4*D2 + 2*D3; (Forma en la que descomponemos a la matriz D.)

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \tag{D}$$

(%i38) P1.P1; (Comprobamos que la matriz de proyección P1 elevada al cuadrado es igual a P1.)

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{\%o38}$$

(%i39) P2.P2; (Comprobamos que la matriz de proyección P2 elevada al cuadrado es igual a P2.)

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{\%o39}$$

(%i40) P3.P3; (Comprobamos que la matriz de proyección P3 elevada al cuadrado es igual a P3.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{\%o40}$$

(%i41) P1T: transpose(P1); (Comprobamos que la matriz de proyección P1 transpuesta es igual a P1.)

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P1T}$$

(%i42) P2T: transpose(P2); (Comprobamos que la matriz de proyección P2 transpuesta es igual a P2.)

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P2T}$$

(%i43) P3T: transpose(P3); (Comprobamos que la matriz de proyección P3 transpuesta es igual a P3.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3T}$$

(%i44) P1.P2; (Al multiplicar las matrices de proyección P1, P2 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o44}$$

(%i45) P1.P3; (Al multiplicar las matrices de proyección P1, P3 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o45}$$

(%i46) P2.P1; (Al multiplicar las matrices de proyección P2, P1 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o46}$$

(%i47) P2.P3; (Al multiplicar las matrices de proyección P2, P3 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o47}$$

(%i48) P3.P1; (Al multiplicar las matrices de proyección P3, P1 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o48}$$

(%i49) P3.P2; (Al multiplicar las matrices de proyección P3, P2 obtenemos la matriz nula.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o49}$$

(%i50) I: P1 + P2 + P3; (Comprobamos que la suma de las matrices de proyección es igual a la matriz identidad $I_{4\times4}$.)

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

(%i51) U1T: matrix((1/2)*[1,1,1,1], [0,0,0,0], [0,0,0,0], [0,0,0,0])$

(%i52) U1T; (Matriz cuyo primer renglón es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o52}$$

(%i53) U1: transpose(U1T); (Matriz cuya primer columna es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} \tag{U1}$$

(%i54) P1: U1.U1T; (Otra forma de obtener la matriz de proyección P1 es multiplicando las matrices formadas por el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 8.)

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P1}$$

(%i55) U2T: matrix([0,0,0,0], (1/2)*[1,-1,-1,1], [0,0,0,0], [0,0,0,0])$

(%i56) U2T; (Matriz cuyo segundo renglón es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{1} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{\%o56}$$

(%i57) U2: transpose(U2T); (Matriz cuya segunda columna es el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix} \tag{U2}$$

(%i58) P2: U2.U2T; (Otra forma de obtener la matriz de proyección P2 es multiplicando las matrices formadas por el vector ortonormal que forma una base ortonormal para el subespacio vectorial correspondiente al valor propio 4.)

$$\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \tag{P2}$$

(%i59) U3T: matrix([0,0,0,0], [0,0,0,0], (1/sqrt(2))*[-1,0,0,1], (1/sqrt(2))*[0,-1,1,0])$

(%i60) U3T; (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \tag{\%o60}$$

(%i61) U3: transpose(U3T); (Matriz formada por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \tag{U3}$$

(%i62) P3: U3.U3T; (Otra forma de obtener la matriz de proyección P3 es multiplicando las matrices formadas por los vectores ortonormales que forman una base ortonormal para el subespacio vectorial correspondiente al valor propio 2.)

$$\begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \tag{P3}$$

(%i63) C: 8*P1 + 4*P2 + 2*P3; (Obtenemos la descomposición espectral de la matriz C.)

$$\begin{pmatrix} 4 & 1 & 1 & 2 \\ 1 & 4 & 2 & 1 \\ 1 & 2 & 4 & 1 \\ 2 & 1 & 1 & 4 \end{pmatrix} \tag{\%C}$$

Acabamos de mostrar que se cumplen todas las propiedades del Teorema de descomposición espectral para la matriz HH^T.

Comentario 3.7. *Una vez revisados los Ejemplos 3.1 y 3.2 es importante mencionar que de acuerdo al tamaño de la matriz* A *el cálculo de los valores propios se vuelve complicado. Porque como el tamaño de la matriz* A *determina el grado del polinomio característico* $P_A(t)$*, entonces mientras mayor sea el tamaño mayor será el grado. Así, encontrar las raíces de tal polinomio se vuelve una tarea no tal fácil. Y a pesar del apoyo que algunos softwares nos ofrecen es indispensable mencionar que, estos tienen una cierta capacidad límite para realizar cálculos con matrices de gran tamaño.*

3.3 ESPECTRO DE UN GRAFO

En esta sección presentamos algunos resultados sobre el espectro de un grafo, en especial de los grafos regulares y los grafos bipartitos. La teoría revisada aquí está enfocada al estudio de grafos finitos, no dirigidos y simples; además, estos resultados nos serán útiles para el estudio de los códigos LDPC.

Los textos revisados para el desarrollo de esta sección fueron [ADH98], [BH12], [Big74], [CDS80], [CRS97], [CRS10] y [MM64].

3.3.1 Espectro de un grafo no dirigido y algunas de sus propiedades

El espectro de un grafo finito $\mathcal{G}$ es por definición el espectro de su matriz de adyacencia A, es decir, el conjunto de los valores propios junto con sus multiplicidades. Como A es una matriz simétrica de valores reales, cada uno de sus valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ son reales, por el Teorema 3.1, y en este caso los valores propios son ordenados de la siguiente manera $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. Además, para cada valor propio la multiplicidad algebraica coincide con la multiplicidad geométrica.

Recordemos que el valor propio más grande de un grafo es conocido como su radio espectral o índice de acuerdo a la Definición 3. La información básica acerca del valor propio más grande de un grafo es proveniente del Teorema de Perron-Frobenius 3.4, por lo cual tenemos la siguiente proposición que se cumple para grafos no dirigidos y posiblemente para digrafos.

Proposición 3.1. *([BH12], pág. 33) Cada grafo* $\mathcal{G}$ *tiene un valor propio real* λ_0 *con su correspondiente vector propio real no negativo y tal que para cada valor propio* λ *tenemos que* $|\lambda| \leqslant \lambda_0$*. El valor* λ_0 *no incrementa cuando los vértices o aristas son removidos de* $\mathcal{G}$*. Asumimos que* $\mathcal{G}$ *es fuertemente conexo. Entonces:*

i) λ_0 *tiene multiplicidad* 1.

ii) Si $\mathcal{G}$ *es primitivo (fuertemente conexo, y tal que no todos los ciclos tienen una longitud que es un múltiplo de algún entero* $m > 1$*), entonces* $|\lambda| < \lambda_0$ *para todos los valores propios* λ *diferentes de* λ_0*.*

iii) El valor λ_0 *disminuye cuando los vértices o aristas son removidos de* $\mathcal{G}$*.*

La siguiente proposición [Big74] está enfocada a grafos k-regulares, y podría decirse que es una versión similar a la Proposición 3.1.

Proposición 3.2. *Sea* $\mathcal{G}$ *un grafo* k*-regular. Entonces:*

i) k *es un valor propio de* $\mathcal{G}$*.*

ii) Si $\mathcal{G}$ *es conexo, entonces la multiplicidad de* k *es uno.*

iii) Para cada valor propio λ *de* $\mathcal{G}$*, tenemos que* $|\lambda| \leqslant k$*.*

Demostración. Sea $\mathcal{G}$ un grafo k-regular.

i) Como $\mathcal{G}$ es un grafo k-regular, entonces cada renglón y cada columna de la matriz de adyacencia A tiene precisamente k $1's$ (recordar que las componentes de la matriz de adyacencia son $0's$ y $1's$). Ahora, sea el vector $u = (1, 1, \ldots, 1)^T$, tal que u satisface $Au = ku$. Así, k es un valor propio de $\mathcal{G}$.

ii) Sea $x = (x_1, x_2, \ldots, x_n)^T$ que denota cualquier vector no cero para el cual $Ax = kx$, y sea x_j una entrada de x que tiene el valor absoluto más grande. Entonces, al multiplicar el j-ésimo renglón de A por el vector x tenemos que $(Ax)_j = kx_j$, además, $(Ax)_j = \sum_{i=1}^{n} x_i$, donde los x_i son las entradas del vector x que quedan cuando las componentes del j-ésimo renglón de la matriz A es diferente de 0, luego, $\sum_{i=1}^{n} x_i = kx_j$, donde la sumatoria es sobre los k vértices que son adyacentes al j-ésimo vértice del grafo. Por la propiedad maximal de x_j, se sigue que $x_i = x_j$ para todos los vértices. Si $\mathcal{G}$ es conexo podemos proceder sucesivamente de este modo, eventualmente demostrando que todas las entradas de x son iguales. Así, x es un múltiplo de u, y el espacio propio asociado con el valor propio k es de dimensión uno.

iii) Supongamos que $Ay = \lambda y$, $y \neq 0$, y sea y_j que denota una entrada de y la cual es la más grande en valor absoluto. Por el mismo argumento como en ii), tenemos que $\sum y_i = \lambda y_j$, y así $|\lambda||y_j| = |\sum y_i| \leqslant k|y_j|$, pues el grafo es regular. Por lo tanto, $|\lambda| \leqslant k$.

□

Observación 3.4. *De la Proposición 3.2, podemos decir que si $\mathcal{G}$ es un grafo k-regular entonces el valor propio más grande es $\lambda_0 = k$. Cuando tenemos un grafo bipartito (l, r)-regular, entonces el valor propio más grande en valor absoluto es $\lambda_0 = l \cdot r$, es decir, es el producto de los grados de sus conjuntos de vértices.*

Proposición 3.3. *Sea $\mathcal{G}$ un grafo con valores propios $k = \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$. Las siguientes afirmaciones son equivalentes:*

i) $\mathcal{G}$ es regular de grado $\lambda_1 = k$.

ii) $AJ = kJ$ donde J es la matriz con todas sus entradas 1, ó $A\vec{1} = k\vec{1}$ donde $\vec{1}$ es el vector con todas sus entradas 1.

iii) $\sum \lambda_i^2 = kn$.

Demostración. Supongamos que $\mathcal{G}$ es un grafo k-regular y sea A su matriz de adyacencia de tamaño $n \times n$ en donde cada renglón y cada columna de A tiene precisamente k $1's$, luego, considerando a J como la matriz con todas sus componentes 1 de tamaño $n \times n$, entonces, al multiplicar cada renglón de A por cada columna de J tenemos que,

$$AJ = \begin{pmatrix} k & k & \cdots & k \\ k & k & \cdots & k \\ \vdots & \vdots & \ddots & \vdots \\ k & k & \cdots & k \end{pmatrix} = k \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} = kJ.$$

En particular, como $\mathcal{G}$ es un grafo k-regular, entonces cada renglón y cada columna de la matriz de adyacencia A tiene precisamente k $1's$. Por lo tanto, el vector $\vec{1} = (1, 1, \ldots, 1)^T$ satisface que $A\vec{1} = k\vec{1}$, así, $\vec{1} = (1, 1, \ldots, 1)^T$ es un vector propio de A correspondiente al valor propio k.

Ahora, si $\mathcal{G}$ es un grafo k-regular recordemos el número de aristas puede ser calculado de la siguiente manera: $|E| = \frac{1}{2}\sum_{i=1}^{n} d(v_i)$, ver Comentario 2.16; pero también tenemos la siguiente relación: $|E| = \frac{1}{2}\sum_{i=1}^{n} \lambda_i^2$, consultar [CDS80] pág. 21. Entonces al igualar las relaciones conseguimos que,

$$\begin{aligned}
& \frac{1}{2}\sum_{i=1}^{n} d(v_i) = \frac{1}{2}\sum_{i=1}^{n} \lambda_i^2 \\
\Rightarrow \quad & \sum_{i=1}^{n} d(v_i) = \sum_{i=1}^{n} \lambda_i^2 \\
\Rightarrow \quad & \sum_{i=1}^{n} d(v_i) \cdot \frac{|V|}{|V|} = \sum_{i=1}^{n} \lambda_i^2 \\
\Rightarrow \quad & d(\mathcal{G})|V| = \sum_{i=1}^{n} \lambda_i^2 \quad \text{(por Definición 2.35 de grado promedio)} \\
\Rightarrow \quad & kn = \sum_{i=1}^{n} \lambda_i^2.
\end{aligned}$$

Por lo tanto, $\sum_{i=1}^{n} \lambda_i^2 = kn$.

Ahora, si se cumple que $\sum_{i=1}^{n} \lambda_i^2 = kn$, entonces $k = \frac{\sum_{i=1}^{n} \lambda_i^2}{n} = \frac{2}{n}\frac{\sum_{i=1}^{n} \lambda_i^2}{2} = \frac{2}{n}|E| = \frac{2}{n}\frac{1}{2}\sum_{i=1}^{n} d(v_i) = \frac{1}{|V|}\sum_{i=1}^{n} d(v_i) = d(G)$. Además, por hipótesis $\lambda_1 = k$, es decir, k es un valor propio del grafo G, por lo que se cumple que $Ax = kx$ en particular con $x = \vec{1}$, es decir, $A\vec{1} = k\vec{1}$, con lo que al multiplicar cada renglón de la matriz de adyacencia A por el vector columna con todas sus entradas 1, tenemos $\sum_{i=1}^{n} 1 = k \cdot 1$, luego, cada renglón de A tiene exactamente k componentes diferentes de cero, con esto, cada vértice del grafo tiene exactamente k vértices adyacentes, y como $d(G) = k$, entonces se satisface que G es un grafo k-regular.

□

Observación 3.5. *Del inciso ii) de la Proposición 3.3 tenemos que $A\vec{1} = k\vec{1}$ con $\vec{1}$ el vector con todas sus entradas 1, entonces el vector $\vec{1}$ es un vector propio de A con su correspondiente valor propio $\lambda_1 = k$ (que es el valor propio más grande). En especial, el vector $v_1 = \frac{1}{\sqrt{n}} \cdot \vec{1} = (1/\sqrt{n}, \ldots, 1/\sqrt{n})$ sigue siendo un vector propio de A correspondiente al mismo valor propio $\lambda_1 = k$, pues $Av_1 = A(\frac{1}{\sqrt{n}} \cdot \vec{1}) = \frac{1}{\sqrt{n}}(A\vec{1}) = \frac{1}{\sqrt{n}}(k\vec{1}) = k(\frac{1}{\sqrt{n}} \cdot \vec{1}) = kv_1$.*

Lema 3.4. *([HLW06], pág. 453) Un grafo G es conexo si y sólo si $\lambda_1 > \lambda_2$ (es decir, λ_1 tiene multiplicidad 1).*

Proposición 3.4. *Si G es un grafo bipartito con al menos una arista, entonces su espectro es simétrico con respecto a 0, es decir, si λ es un valor propio de G, entonces $-\lambda$ es también un valor propio de G, con la misma multiplicidad.*

Demostración. Sea G un grafo bipartito con matriz de adyacencia de la forma $A = \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix}$. Supóngase que λ es un valor propio de G y $x = (x_1, x_2, \ldots, x_n)$ es un vector propio correspondiente. Considere el vector $y = (y_1, y_2, \ldots, y_n)$ donde $y_j = x_j$ si $1 \leqslant j \leqslant m$ y $y_j = -x_j$ si $m+1 \leqslant j \leqslant n$. Entonces,

$$\sum_{j=1}^{n} a_{ij}y_j = \sum_{j=m+1}^{n} a_{ij}y_j = -\sum_{j=m+1}^{n} a_{ij}x_j = -\lambda x_i = -\lambda y_i, \quad \text{si} \quad 1 \leqslant i \leqslant m,$$

y

$$\sum_{j=1}^{n} a_{ij}y_j = \sum_{j=1}^{m} a_{ij}y_j = \sum_{j=1}^{m} a_{ij}x_j = \lambda x_i = -\lambda y_i, \quad \text{si} \quad m+1 \leqslant i \leqslant n.$$

Así, y satisface que $Ay = -\lambda y$, y $-\lambda$ es un valor propio de G.

□

H. Sachs (1966) demostró que la simetría del espectro de G es suficiente para que G sea bipartito. Con lo cual se tiene el siguiente teorema.

Teorema 3.13. *Un grafo G con al menos una arista es bipartito si y sólo si su espectro es simétrico con respecto a 0.*

Incluimos una caracterización fuerte de los grafos bipartitos regulares realizada por Hoffman en 1963.

Teorema 3.14. *Un grafo G conexo k-regular es bipartito si y sólo si $-k$ es un valor propio de G.*

Demostración. Primero supóngase que G es un grafo bipartito k-regular. Entonces por el inciso ii) de la Proposición 3.3 tenemos que el vector $\vec{1} = (1, 1, \ldots, 1)^T$ satisface que $A\vec{1} = k\vec{1}$, con k un valor propio de G y así, por la Proposición 3.4, $-k$ es también un valor propio de G.

Sea $G = (V, E)$ un grafo k-regular con su conjunto de vértices $V = \{v_1, v_2, \ldots, v_n\}$. Para demostrar el recíproco debemos construir una bipartición de G. Sea $x = (x_1, x_2, \ldots, x_n)^T$ un vector propio de G, con valor propio $-k$, es decir, $Ax = -kx$. Podemos asumir, sin pérdida de generalidad que, $\max_{1 \leqslant j \leqslant n} x_j = 1$ y que $|x_j| \leqslant 1$ para cada j.

Supongamos que $x_{i_0} = 1$. Luego, como cada renglón y columna de A contiene precisamente k $1's$ y $\sum_{j=1}^{n} a_{i_0 j} x_j = -k x_{i_0} = -k$, entonces x_j debería ser igual a -1 para cada j tal que $v_j \in N(v_{i_0})$ donde $N(v_{i_0})$ es el conjunto de vecinos del vértice v_{i_0}. Similarmente, tenemos que si $x_{i_0} = -1$ entonces $x_j = 1$ para cada j con $v_j \in N(v_{i_0})$. Como $\mathcal{G}$ es conexo todas las coordenadas de x deben ser, por lo tanto, 1 ó -1. Sea $V_1 = \{v_j : x_j = 1, 1 \leqslant j \leqslant n\}$ y $V_2 = \{v_j : x_j = -1, 1 \leqslant j \leqslant n\}$. Entonces (V_1, V_2) es una bipartición de V, así, $\mathcal{G}$ es un grafo bipartito. □

Observemos la siguiente relación que existe entre los valores propios de la matriz A y HH^T cuando A es una matriz de adyacencia del grafo $\mathcal{G}$ de la forma $A = \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix}$.

Lema 3.5. *Sea A una matriz de adyacencia de un grafo bipartito regular conexo. Los valores propios de A son $\pm\sqrt{\mu_i}$ y posiblemente 0, donde μ_i, $1 \leqslant i \leqslant s$ son los distintos valores propios de HH^T.*

Demostración. Sea $\mathcal{G}$ un grafo bipartito con matriz de adyacencia $A \in M(\mathbb{F}_2)_{(m+n)\times(m+n)}$ definida como

$$A = \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix}$$

notemos que A es una matriz por bloques, donde el primer bloque 0 es la matriz nula de tamaño $m \times m$, el segundo bloque H es una matriz de tamaño $m \times n$, el tercer bloque H^T es la matriz transpuesta de H de tamaño $n \times m$, y el cuarto bloque 0 es la matriz nula de tamaño $n \times n$.

Entonces,

$$\begin{aligned} A^2 &= A \cdot A \\ &= \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix} \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} HH^T & 0 \\ 0 & H^T H \end{bmatrix} \end{aligned}$$

donde el primer bloque HH^T es una matriz de tamaño $m \times m$, el segundo bloque 0 es la matriz nula de tamaño $m \times n$, el tercer bloque 0 es la matriz cero de tamaño $(n \times m$, y el cuarto bloque $H^T H$ es la matriz de tamaño $n \times n$.

Ahora bien, calculemos el polinomio característico de A^2,

$$\begin{aligned} P_{A^2}(t) &= \det(tI - A^2) \\ &= \left|tI - A^2\right| \\ &= \left|tI - \begin{bmatrix} HH^T & 0 \\ 0 & H^T H \end{bmatrix}\right| \\ &= \begin{vmatrix} tI - HH^T & 0 \\ 0 & tI - H^T H \end{vmatrix} \\ &= \left|tI - HH^T\right| \cdot \left|tI - H^T H\right| \\ &= P_{HH^T}(t) \cdot P_{H^T H}(t) \end{aligned} \tag{3.17}$$

Usando el Lema 3.1 obtenemos que $P_{H^T H}(t) = t^k P_{HH^T}(t)$ con $k = n - m$, entonces de (3.17) $P_{A^2}(t) = P_{HH^T}(t) \cdot P_{H^T H}(t) = P_{HH^T}(t) \cdot t^k P_{HH^T}(t) = t^k P_{HH^T}(t) \cdot P_{HH^T}(t) = t^k (P_{HH^T}(t))^2$, eso implica que $P_{A^2}(t) = t^k (P_{HH^T}(t))^2$. Al despejar a P_{HH^T} tenemos que $P_{HH^T}(t) = \pm\sqrt{t^{-k} P_{A^2}(t)}$.

Consideremos a $\mu_i \neq 0$ con $1 \leqslant i \leqslant s$ los distintos valores propios de HH^T, entonces para cada i,

$$\begin{aligned} & P_{H^T H}(\mu_i) = 0 \\ \Rightarrow \quad & \sqrt{\mu_i^{-k} P_{A^2(\mu_i)}} = 0 \\ \Rightarrow \quad & \mu_i^{-k} = 0 \quad \vee \quad P_{A^2}(\mu_i) = 0, \end{aligned}$$

con lo cual, si se satisface que $P_{A^2(\mu_i)} = 0$ entonces obtendríamos que μ_i también es un valor propio de la matriz A^2, y usando el Lema 3.2 obtenemos que $\pm\sqrt{\mu_i}$ son los valores propios de A. □

4 CÓDIGOS LDPC

El razonamiento matemático puede considerarse
más bien esquemáticamente como el ejercicio
de una combinación de dos instalaciones,
que podemos llamar la intuición y el ingenio.

– Alan Turing –

En este capítulo presentamos los códigos LDPC, damos un pequeño acercamiento al estudio de su teoría fundamental. Una vez revisados los capítulos previos que tienen que ver en parte con la teoría de códigos, la teoría de grafos (o gráficas), la teoría de la información, y el álgebra lineal; nos es posible realizar un estudio sobre estos códigos que permiten comunicar con eficiencia y confiabilidad la información. No olvidemos que el estudio de los códigos LDPC se debe a que son una excelente familia de códigos detectores-correctores de errores que intentan dar solución al problema central de la teoría de comunicación, cuyo objetivo es construir un sistema de codificación y decodificación que haga posible una casi perfecta comunicación sobre el canal con ruido.

El objetivo central de este capítulo es estudiar el concepto de código LDPC regular e irregular, considerando su matriz de chequeo de paridad, como la matriz relacionada con un sistema de ecuaciones lineales, y también, como la matriz de adyacencia de un grafo bipartito. Luego, caracterizamos la irregularidad de un código LDPC a partir del cálculo de las distribuciones de grado. Y realizaremos un análisis detallado sobre las demostraciones de tres teoremas que establecen una cota para la distancia mínima del código. Además, damos algunos ejemplos de la teoría revisada usando el software *Magma V2.14-15* [CBFS08]. No sin antes hacer una introducción, dando un panorama general de lo que abarca el estudio, la implementación y las aplicaciones de los códigos LDPC.

Los textos revisados para el desarrollo de este último capítulo fueron [Gal63], [Kel09], [RU08], [Tan01] y [WK03].

4.1 UNA INTRODUCCIÓN A LOS CÓDIGOS LDPC

En esta sección hablaremos sobre cómo y quién introdujó el concepto de los ***códigos de Chequeo de Paridad de Baja Densidad*** conocidos como códigos **LDPC** por sus siglas en inglés ***Low Density Parity Check***. Introduciremos que es lo que caracteriza a esta clase de códigos, hablaremos un poco de su codificación y decodificación, y mencionaremos algunas de las aplicaciones donde estos códigos LDPC posiblemente están en uso. Cabe señalar, que lo mencionado en esta sección no necesariamente será abordado con más a detalle, debido a que el trabajo de los códigos LDPC es muy extenso, y se requiere de un amplio conocimiento sobre diversas áreas de las matemáticas para llevar a cabo el análisis de este panorama general.

4.1.1 Historia

Robert G. Gallager desarrolló el concepto de códigos LDPC en su tesis doctoral en el MIT en 1960 y después fue publicada en 1963, debido a esto, los códigos LDPC también son conocidos como códigos de Gallager en su honor. En su trabajo Gallager introdujo la definición de códigos de chequeo de paridad de baja densidad y también la decodificación iterativa, [Gal63]. Sin embargo, los códigos LDPC fueron olvidados por la alta complejidad computacional que suponía su decodificación con la tecnología de entonces. Pinsker y Zyablov fueron los primeros que analizaron la decodificación iterativa. En 1981 Robert M. Tanner generalizó la construcción de Gallager y estableció una representación gráfica de los códigos en términos de grafos bipartitos, [Tan01]. A mitad de los años 90 cuando se introduce el

concepto de turbo códigos, David Mackay y Neal redescubrieron los códigos de Gallager, siendo más viable su implementación para ese entonces. Hay varios trabajos que involucran a los códigos LDPC y que se relacionan con todo el trabajo realizado por los personajes mencionados y los no mencionados, como el análisis de cotas para la distancia mínima, [Tan01] y [SKS05]; posteriormente, usando el concepto de expansión de un grafo se construye una nueva familia de códigos, los códigos expandidos construidos por Sipser y Spielman que son un caso especial de la clase de códigos LDPC, [SS96]; y de igual forma el análisis sobre su distancia mínima, [FZ11]. También se han investigado las distribuciones de peso y grado de los códigos LDPC, entre algunos autores tenemos a Di, Richardson, y Urbanke; una idea sugerida por Chung, Richardson, y Urbanke fue obtener una aproximación Gaussiana como una aproximación unidimensional de la evolución de densidad para los códigos LDPC, [RU01] y [RU08].

Hay más trabajos realizados y relacionados con el estudio de los códigos LDPC, así que con el párrafo anterior se pretende que el lector conozca de manera general lo que a lo largo del tiempo se ha hecho. Hoy en día los códigos LDPC están en gran auge, el interés por estos sistemas de codificación y decodificación basados en estos códigos, se debe a que pueden aproximarse bastante bien al límite de capacidad establecido en el Teorema de Shannon 2.8.

4.1.2 Notación

Descripción de la notación implementada.

Notación	Descripción
$\mathcal{C}(\mathcal{B})$	Código LDPC.
$\mathcal{B}$	Grafo bipartito.
X	Conjunto de vértices o nodos variable.
C	Conjunto de vértices o nodos restricción.
x_i	Nodo variable o nodo palabra-código o nodo bit.
c_i	Nodo restricción o nodo redundancia o nodo chequeo de paridad.
n	Número de nodos bit.
m	Número de nodos chequeo de paridad.
d_{x_i}	Grado del i-ésimo nodo variable.
d_{c_i}	Grado del i-ésimo nodo restricción.
d_x	Grado de los nodos variable cuando el grafo bipartito es regular o grado máximo de los nodos variable cuando el grafo bipartito es irregular.
d_c	Grado de los nodos restricción cuando el grafo bipartito es regular o grado máximo de los nodos restricción cuando el grafo bipartito es irregular.
a_x	Grado promedio de los nodos variable.
a_c	Grado promedio de los nodos restricción.
w_{x_i}	Peso de la i-ésima columna de la matriz de chequeo de paridad.
w_{c_i}	Peso del i-ésimo renglón de la matriz de chequeo de paridad.
w_x	Peso mínimo de las columnas o peso de las columnas si todas tienen el mismo peso.
w_c	Peso mínimo de los renglones o peso de los renglones si todos tienen el mismo peso.
d	Distancia mínima.
k	Dimensión del código.
R	Tasa del código.
M	Número de palabras código.
H	Matriz de chequeo de paridad del código de tamaño $(n-k) \times n$.
G	Matriz generadora del código de tamaño $k \times n$.
A	Matriz de adyacencia del grafo bipartito de tamaño $n \times n$.

4.1.3 Nociones y características de los códigos LDPC

Los códigos LDPC tienen una estructura teoríca muy interesante a continuación mencionaremos algunos puntos esenciales.

- Los códigos LDPC son una clase de códigos de bloque lineal cuya propiedad esencial es la de tener una matriz de chequeo de paridad H dispersa o de baja densidad, es decir, con pocos elementos distinto de cero.
- Pueden ser tratados como un grafo bipartito $\mathcal{B} = (X \cup C, E)$ cuya matriz de adyacencia es de la forma $A = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix}$, donde H es la matriz de chequeo de paridad del código LDPC. En el grafo bipartito $\mathcal{B}$, X representa la clase de n nodos conocidos como nodos palabras-código, variables o bits que los denotaremos como x; y C representa la clase de m nodos conocidos como nodos restricción, redundancia o chequeo que los denotaremos como c. Esta representación es conocida como el grafo de Tanner.
- La condición impuesta por los nodos chequeo c es que la suma de sus vecinos, es decir, los nodos variable x deben sumar 0 módulo 2.
- Los códigos LDPC pueden ser regulares o irregulares.
- El hecho de poder representar los códigos LDPC mediante grafos bipartitos, ayuda a entender la estructura del código y proporciona un poderoso enfoque de decodificación, ya que los algoritmos eficientes de decodificación se basan en esta representación.
- La decodificación eficiente de los códigos LDPC se logra mediante los algortimos que se engloban en los llamados algoritmos de paso de mensaje, son algoritmos iterativos y su nombre se debe a que el algoritmo de decodificación se basa en la observación de la estructura del grafo bipartito y en como en cada iteración, se pasan mensajes de los nodos bit a los nodos chequeo y viceversa.
- Una importante subclase de los algoritmos de paso de mensaje son los llamados algoritmos Belief-Propagation (BP) que fueron presentados por Gallager. Entre los cuales, existen diversos algoritmos de decodificación de códigos LDPC basados en los BP como el Suma-Producto (SP), y la simplificación Min-Sum (MS).
- El estudio de esta clase de códigos comenzó con los códigos LDPC binarios pero ya existen trabajos que estudian a los códigos LDPC no binarios.
- También se conoce el código dual para los códigos LDPC llamados códigos de Matriz Generadora de Baja Densidad conocidos como LDGM, en su versión en inglés Low Density Generator Matrix.
- Progresivamente, surgen los códigos cuasi-cíclicos de los códigos LDPC conocidos como (QC)-LDPC, teniendo una estructura interior que le permite considerablemente hacer más pequeño el software y complejidad de implementación del hardware.
- Un caso especial son los códigos expandidos que pertenecen a la clase de los códigos LDPC, de la cual podemos decir que, si el grafo bipartito $\mathcal{B}$ es una buena expansión y si los nodos restricción vecinos a un subconjunto de los nodos variable son identificados como subcódigos suficientemente buenos, entonces el código expandido resultante será un buen código, y por tanto, tendrá buena propiedad de distancia mínima.

Como uno de los objetivos es proporcionar ese acercamiento a los códigos LDPC los puntos anteriores fueron mencionados, sin embargo, no todo lo enunciado con anterioridad será abordado, ya que nuestro interés va más enfocado a obtener cotas para la distancia mínima de dichos códigos.

4.1.4 Algunas de las aplicaciones posibles de los códigos LDPC

Los códigos LDPC se están implementando en aplicaciones donde la transferencia de información a través del ancho de banda o de canal de retorno está limitado por la presencia de ruido. Esto es, está siendo incluido en multitud de estándares de comunicación debido a su gran capacidad de detección y corrección, así como su facilidad para paralelizar el proceso de decodificación, lo que permite su uso en sistemas que requieren altas velocidades de transmisión. Algunas de sus posibles implementaciones son, [Lhé04]:

- Recuperación de paquetes perdidos en la distribución de datos masivos a varios clientes en forma simultánea a través de la Internet.
- Almacenamiento en medios magnéticos.
- Almacenamiento distribuido de información.
- Corrección de errores en telefonía común e inalámbrica y en módems.
- En el año 2003, un código LDPC venció a 6 códigos turbo para convertirse en la corrección de errores de códigos en el nuevo estándar DVB-S2 para la transmisión por satélite de televisión digital.
- En el año 2008, los códigos LDPC fueron escogidos como códigos para ser utilizados en el régimen FEC para el UIT-T. Los códigos LDPC también se utiliza para 10GBASE-T Ethernet a través de cables categoría CAT6.

4.2 LOS CÓDIGOS LDPC

En la sección anterior 4.1 dimos una idea general de los códigos de *chequeo de paridad de baja densidad*. Ahora, en esta sección estudiaremos formalmente el concepto de los códigos *LDPC*, analizando su sistema de ecuaciones lineales y su representación gráfica, y viendo como estas dos ideas se relacionan con su matriz de chequeo de paridad. Y además, daremos un par de ejemplos sobre un código LDPC regular e irregular.

El campo sobre el que trabajamos es el campo binario $\mathbb{F}_2$, entonces nos enfocamos en el estudio de los códigos LDPC binarios.

4.2.1 Definición de un código LDPC y su representación gráfica

De acuerdo al trabajo de Gallager [Sha48] los códigos LDPC se definen como.

Definición 4.1. *Los códigos LDPC son códigos de bloque lineal determinados por una matriz* H *cuyas componentes en su mayor parte son* $0's$ *y relativamente pocos* $1's$*, es decir, una matriz de chequeo de paridad* H *dispersa o de baja densidad.*

Observación 4.1. *Un código LDPC regular es un código de chequeo de paridad de baja densidad en el cual cada columna de* H *tiene el mismo peso* w_x *y cada renglón de* H *tiene el mismo peso* w_c*. Un código LDPC irregular es un código en el cual las columnas y/o los renglones de* H *no tienen el mismo peso.*

En la sección 2.5 del capítulo 2 revisamos la relación que existe entre un código lineal y los grafos bipartitos. Entonces, debido a que los códigos LDPC son códigos lineales, estos tienen una representación gráfica conocida como grafo de Tanner. El grafo resultante del código LDPC será un grafo bipartito, y además, su matriz de adyacencia A está relacionada con la matriz de chequeo de paridad H, ya que $A = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix}$. Si el grafo bipartito es regular decimos que obtenemos un código LDPC regular, y si el grafo bipartito es irregular entonces obtenemos un código LDPC irregular, [Tan01] y [RU08].

Una vez relacionada la definición dada por Gallager con la representación gráfica de Tanner, se tiene la siguiente definición de un código LDPC, [WK03].

Definición 4.2. *Sea* $\mathcal{B} = (X \cup C, E)$ *un grafo bipartito donde* X *es el conjunto de* n *nodos variables* $x_1, x_2, \ldots, x_n$*,* C *es el conjunto de los* m *nodos restricción* $c_1, c_2, \ldots, c_m$*, y* E *es el conjunto de aristas. Un código* $\mathcal{C}(\mathcal{B})$ *de chequeo de paridad de baja densidad o simplemente un código LDPC de longitud* n *es*

$$\mathcal{C}(\mathcal{B}) = \left\{ (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mid x_{b(1,j)} \oplus x_{b(2,j)} \oplus \cdots \oplus x_{b(d_{c_j},j)} = 0, 1 \leqslant j \leqslant m \right\}, \tag{4.1}$$

donde $b(i,j)$ *es una función incidente definida tal que para cada nodo restricción* c_j*, los nodos vecinos o colindantes a* c_j *son* $x_{b(1,j)}$*,* $x_{b(2,j)}$*, ...,* $x_{b(d_{c_j},j)}$*, y los grados de cada nodo restricción* $d_{c_{j's}}$ *son muy pequeños comparados con el número de los* n *nodos variables.*

Observación 4.2. *La función incidente asocia un par no ordenado de distintos nodos con cada arista del grafo bipartito $\mathcal{B}$, es decir, $\{x_{b(1,j)}, c_j\}$, $\{x_{b(2,j)}, c_j\}$, ..., $\{x_{b(d_{c_j},j)}, c_j\}$ forman las aristas de $\mathcal{B}$.*

Observación 4.3. *A partir de la Definición 4.2 podemos decir que un código LDPC, es el conjunto de las soluciones de un sistema de m ecuaciones lineales, donde $\oplus$ indica la suma módulo 2 al estar trabajando en el campo binario $\mathbb{F}_2$. Si a_x y a_c son los grados promedios de los nodos variable y restricción en el grafo $\mathcal{B}$, respectivamente, entonces $m = \frac{a_x \cdot n}{a_c}$.*

Comentario 4.1. *Cabe mencionar que la Definición 4.2 abarca los códigos LDPC regulares e irregulares, pero en los resultados dados posteriormente aclararemos con que tipo de código trabajaremos y daremos los parámetros apropiados.*

Estudiemos el concepto de un código LDPC. Hay dos formas de construir un código LDPC: la primera, a partir de un sistema de ecuaciones lineales, con la cual obtendríamos su matriz de chequeo de paridad H y solo resta verificar que se cumpla que el número de entradas en su mayor parte sean $0's$ y relativamente pocos $1's$, con esto conseguimos un código LDPC. Luego, se sabe que este código lineal tiene una representación gráfica, pero con la matriz H es suficiente para construir el grafo bipartito correspondiente y además obtener la matriz de adyacencia A. La segunda, a partir de tener un grafo bipartito, uno puede obtener su matriz de adyacencia del grafo, y llevarlo a la forma $A = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix}$, de donde obtenemos H la matriz de chequeo de paridad de algún código, así, al verificar que se satisface que el número de de entradas en su mayor parte sean $0's$ y relativamente pocos $1's$, entonces conseguimos un código LDPC.

Observamos que el vínculo entre una u otra forma es analizar la matriz de chequeo de paridad H, por lo que, podemos construir esta matriz H como una matriz aleatoria o una matriz determinista y posteriormente ver si satisface la definición de código LDPC, y entonces decir que el código que tenemos es un código LDPC.

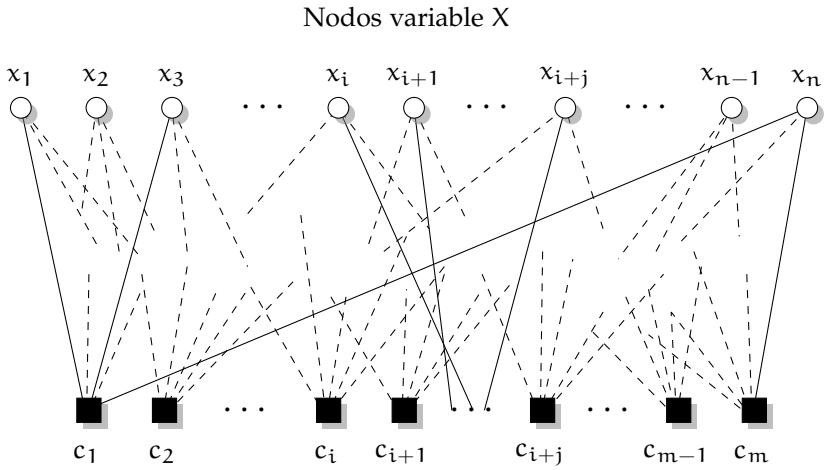
A partir de la Definición 4.2, un código LDPC es el conjunto $\mathcal{C}(\mathcal{B})$ (4.1) donde $\mathcal{B} = (X \cup C, E)$ es el grafo biparito asociado al código y $b(i,j)$ es una función incidente, entonces,

$$\begin{aligned}
\mathcal{C}(\mathcal{B}) &= \left\{(x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mid x_{b(1,j)} \oplus x_{b(2,j)} \oplus \cdots \oplus x_{b(d_{c_j},j)} = 0, 1 \leqslant j \leqslant m\right\} \\
&= \left\{(x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mid \begin{array}{ccccccccccc}
x_{b(1,1)} & \oplus & x_{b(2,1)} & \oplus & \cdots & \oplus & x_{b(d_{c_1},1)} & = & 0 \\
x_{b(1,2)} & \oplus & x_{b(2,2)} & \oplus & \cdots & \oplus & x_{b(d_{c_2},2)} & = & 0 \\
\cdots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots \\
x_{b(1,i)} & \oplus & x_{b(2,i)} & \oplus & \cdots & \oplus & x_{b(d_{c_i},i)} & = & 0 \\
\cdots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \cdots \\
x_{b(1,m)} & \oplus & x_{b(2,m)} & \oplus & \cdots & \oplus & x_{b(d_{c_m},m)} & = & 0
\end{array}\right\}
\end{aligned}$$

Es decir, obtenemos un sistema de m ecuaciones lineales, donde la i-ésima ecuación lineal va ser la suma de los nodos variable adyacentes al i-ésimo nodo restricción c_i, en sí, cada $x_{b(i,j)}$ es algún nodo variable. Luego, la matriz asociada con este sistema de ecuaciones es una matriz con componentes $0's$ y $1's$, a la cual denotamos por H. Así, $\mathcal{C}(\mathcal{B}) = \left\{x \in \mathbb{F}_2^n \mid Hx^T = 0\right\}$ con H una matriz de chequeo de paridad de tamaño $m \times n$, a partir de esto decimos que tenemos un código lineal con palabras-código x de longitud n; además, al ser un código de bloque tenemos la clase de n nodos variable x_i y la clase de m nodos restricción c_i. Ahora, al contar cuantos $1's$ hay en cada renglón de la matriz H obtenemos el peso de cada renglón, esto es, w_{c_i}, y al contar cuantos $1's$ hay en cada columna de la matriz H obtenemos el peso de cada columna, es decir, w_{x_i}.

Después, al contar con la matriz H del código podemos construir el grafo correspondiente, el cual resulta ser un grafo bipartito $\mathcal{B} = (X \cup C, E)$ donde $X = \{x_1, x_2, \ldots, x_n\}$ es el conjunto de los nodos variable que son determinadas por las columnas de la matriz H, $C = \{c_1, c_2, \ldots, c_m\}$ es el conjunto de los nodos restricción que son determinadas por los renglones de la matriz H, y E es el conjunto de las aristas determinadas por las entradas diferentes de cero de la matriz H. Este grafo bipartito puede ser dibujado de alguna de las dos maneras como se ilustra en la Figura 45, o quizás de alguna otra

forma diferente, de acuerdo a como resulte de mayor utilidad. También es posible construir su matriz de adyacencia A del grafo bipartito $\mathcal{B}$.



$$\mathcal{B} = (X \cup C, E)$$

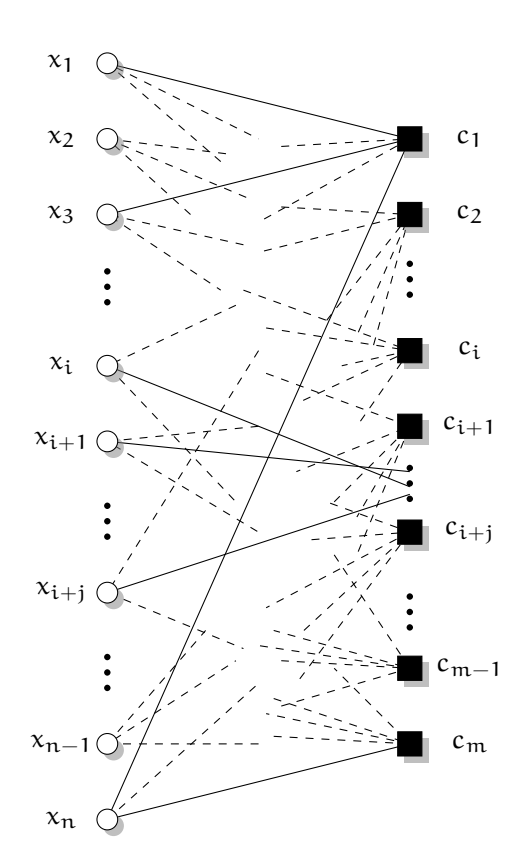


Figura 45: Grafo de Tanner asociado a un código LDPC.

Cuando dibujamos el grafo bipartito $\mathcal{B}$ podemos ver si éste es regular o no, aunque esto también se puede determinar desde la matriz de chequeo de paridad H, pues para cada i el peso de la i-ésima columnas w_{x_i} es igual al grado d_{x_i} del i-ésimo nodo variable y el peso del i-ésimo renglón w_{c_i} resultan ser igual al grado d_{c_i} del i-ésimo nodo restricción. Así, cuando los pesos de las columnas w_{x_i} sean

todos iguales y también los pesos de los renglones w_{c_i} sean todos iguales, aunque no necesariamente pesos iguales entre columna y renglón, diremos que tenemos un grafo bipartito (d_x, d_c)-regular, en caso contrario, será un grafo bipartito irregular con d_x y d_c los grados máximos correspondientes a los nodos variable y restricción, respectivamente.

Un código LDPC al ser un código lineal y contar con una representación gráfica bipartita sus características y propiedades están determinadas por dichas estructuras, y debido a esto es posible obtener resultados interesantes.

4.2.2 Ejemplos de códigos LDPC

Veamos el ejemplo de un código LDPC regular.

Ejemplo 4.1. *Sea* H *la matriz de chequeo de paridad de un código, verifiquemos que el código resultante es un código LDPC* $\mathcal{C}(\mathcal{B})$ *y analicemos que características, propiedades y parámetros tiene.*

Dada la matriz H (4.2) *de un código lineal con la clase de* $n = 8$ *nodos variable* $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ *y la clase de* $m = 4$ *nodos restricción* $C = \{c_1, c_2, c_3, c_4\}$. *Lo primero que observamos de la matriz de chequeo de paridad* H *es que el número de unos es igual al número de ceros, además, que el peso de todos los renglones es* $w_c = 4$ *y el peso de todas las columnas es* $w_x = 2$.

$$H = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{array} \begin{array}{c} \begin{array}{cccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array} \tag{4.2}$$

Luego, obtenemos el sistema de ecuaciones correspondiente a $Hx^T = 0$ *donde* $x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ *es una palabra-código,*

$$x_2 + x_4 + x_5 + x_8 = 0 \tag{4.3}$$
$$x_1 + x_2 + x_3 + x_6 = 0 \tag{4.4}$$
$$x_3 + x_6 + x_7 + x_8 = 0 \tag{4.5}$$
$$x_1 + x_4 + x_5 + x_7 = 0 \tag{4.6}$$

del sistema de ecuaciones la primer ecuación (4.3) *es asociada al primer nodo restricción* c_1, *la segunda ecuación* (4.4) *al segundo nodo restricción* c_2, *la tercer ecuación* (4.5) *al tercer nodo restricción* c_3, *y la cuarta ecuación* (4.6) *al cuarto nodo restricción* c_4.

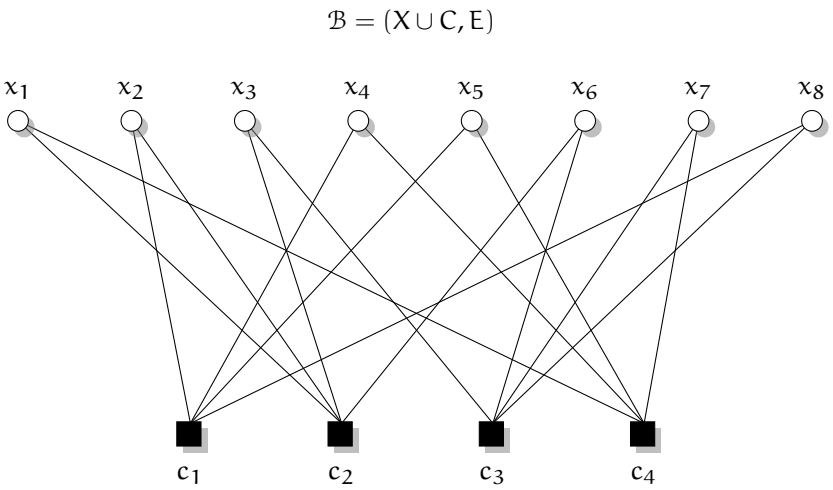


Figura 46: Grafo bipartito regular asociado con un código LDPC regular con matriz de chequeo de paridad H.

Además, contruimos el grafo bipartito $\mathcal{B} = (X \cup C, E)$ *correspondiente a la matriz* H*, que es* $(2,4)$*-regular y donde* $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ *es el conjunto con* 8 *nodos variable y* $C = \{c_1, c_2, c_3, c_4\}$ *es el conjunto con* 4 *nodos restricción, el cual ilustramos en la Figura 46. Al observar el grafo bipartito* $\mathcal{B}$ *vemos que los nodos adyacentes a cada nodo restricción son justamente los nodos variable asociados con las ecuaciones descritas anteriormente.*

Con ayuda del software ***Magma V2.14-15*** *[CBFS08] analicemos la matriz* H*.*

```
> M := KMatrixSpace(FiniteField(2),4,8);
> M1 := M ! [ 0,1,0,1,1,0,0,1,
              1,1,1,0,0,1,0,0,
              0,0,1,0,0,1,1,1,
              1,0,0,1,1,0,1,0 ];
> H := SparseMatrix(M1);
> H;

Sparse matrix with 4 rows and 8 columns over GF(2)

> C := LDPCCode(H);
> C;

[8, 5, 2] Linear Code over GF(2)
Generator matrix:
[1 0 0 0 0 1 1 0]
[0 1 0 0 0 1 0 1]
[0 0 1 0 0 1 0 0]
[0 0 0 1 0 0 1 1]
[0 0 0 0 1 0 1 1]

> CodeWordsC := [x : x in C];
> CodeWordsC;

[
    (0 0 0 0 0 0 0 0),
    (1 0 0 0 0 1 1 0),
    (1 1 0 0 0 0 1 1),
    (0 1 0 0 0 1 0 1),
    (0 1 1 0 0 0 0 1),
    (1 1 1 0 0 1 1 1),
    (1 0 1 0 0 0 1 0),
    (0 0 1 0 0 1 0 0),
    (0 0 1 1 0 1 1 1),
    (1 0 1 1 0 0 0 1),
    (1 1 1 1 0 1 0 0),
    (0 1 1 1 0 0 1 0),
    (0 1 0 1 0 1 1 0),
    (1 1 0 1 0 0 0 0),
    (1 0 0 1 0 1 0 1),
    (0 0 0 1 0 0 1 1),
    (0 0 0 1 1 0 0 0),
    (1 0 0 1 1 1 1 0),
    (1 1 0 1 1 0 1 1),
    (0 1 0 1 1 1 0 1),
    (0 1 1 1 1 0 0 1),
    (1 1 1 1 1 1 1 1),
    (1 0 1 1 1 0 1 0),
    (0 0 1 1 1 1 0 0),
    (0 0 1 0 1 1 1 1),
    (1 0 1 0 1 0 0 1),
```

```
    (1 1 1 0 1 1 0 0),
    (0 1 1 0 1 0 1 0),
    (0 1 0 0 1 1 1 0),
    (1 1 0 0 1 0 0 0),
    (1 0 0 0 1 1 0 1),
    (0 0 0 0 1 0 1 1)
]

> #C;

32

> IsLDPC(C);

true

> IsRegularLDPC(C);

true 4 2

> TannerGraph(C);

Graph
Vertex  Neighbours

1       10 12 ;
2       9 10 ;
3       10 11 ;
4       9 12 ;
5       9 12 ;
6       10 11 ;
7       11 12 ;
8       9 11 ;
9       2 4 5 8 ;
10      1 2 3 6 ;
11      3 6 7 8 ;
12      1 4 5 7 ;
```

Entonces, una vez realizado el análisis con Magma, los resultados obtenidos son que el código $\mathcal{C}(\mathcal{B})$ *es un código LDPC de dimensión* $\dim \mathcal{C}(\mathcal{B}) = 5$*, con palabras-código* x *de longitud* 8*, con distancia mínima* $d = 2$*, y, con pesos en las columnas* $w_x = 2$ *y pesos en los renglones* $w_c = 4$*. Es decir, un* $[8,5,2]$*-código LDPC* $(2,4)$*-regular con tasa* $R = \frac{5}{8}$*, cuya matriz generadora es,*

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

con cardinalidad de $|\mathcal{C}(\mathcal{B})| = 32$ *palabras-código* x *que están enlistadas anteriormente. Además, al ser un código LDPC* $(2,4)$*-regular el grafo bipartito asociado es un grafo bipartito* $\mathcal{B}$ $(2,4)$*-regular, y en la parte final del análisis se obtienen los vecinos de los* 12 *nodos en total del grafo bipartito* $\mathcal{B}$ *y que son justamente los que se observan en la Figura 46. Por lo tanto, el* $[8,5,2]$*-código LDPC con matriz de chequeo de paridad* H *(4.2) y con representación gráfica bipartita* $\mathcal{B}$ *46, es el conjunto*

$$\mathcal{C}(\mathcal{B}) = \left\{ (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \mid x_{b(1,j)} \oplus x_{b(2,j)} \oplus \cdots \oplus x_{b(d_{c_j},j)} = 0, 1 \leqslant j \leqslant 4 \right\}.$$

Veamos el ejemplo de un código LDPC irregular.

Ejemplo 4.2. *Sea* H *la matriz de chequeo de paridad de un código, verifiquemos que el código resultante es un código LDPC,* $\mathcal{C}(\mathcal{B})$*, y analicemos que características, propiedades y parámetros tiene.*

Dada la matriz H (4.7) *de un código lineal con la clase de* $n = 9$ *nodos variable* $X = \{x_1, x_2, x_3, x_4, x_5, x_6 x_7, x_8, x_9\}$ *y la clase de* $m = 5$ *nodos restricción* $C = \{c_1, c_2, c_3, c_4, c_5\}$.

$$H = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{array} \begin{array}{c} \begin{array}{ccccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array} \tag{4.7}$$

Observemos que el número de unos es menor al número de ceros en la matriz H*, además, que el peso de los renglones no es el mismo y tampoco el peso de las columnas es igual. Luego, el sistema de ecuaciones correspondiente a* $Hx^T = 0$ *donde* $x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9)$ *es una palabra-código,*

$$x_2 + x_4 + x_5 + x_6 + x_8 = 0 \tag{4.8}$$
$$x_2 + x_4 + x_6 + x_8 + x_9 = 0 \tag{4.9}$$
$$x_3 + x_4 + x_7 = 0 \tag{4.10}$$
$$x_1 + x_2 + x_5 + x_6 + x_8 = 0 \tag{4.11}$$
$$x_1 + x_4 + x_6 + x_9 = 0 \tag{4.12}$$

del sistema de ecuaciones la primer ecuación (4.8) *es asociada al primer nodo restricción* c_1*, la segunda ecuación* (4.9) *al segundo nodo restricción* c_2*, la tercer ecuación* (4.10) *al tercer nodo restricción* c_3*, la cuarta ecuación* (4.11) *al cuarto nodo restricción* c_4*, y la quinta ecuación* (4.12) *al quinto nodo restricción* c_5.

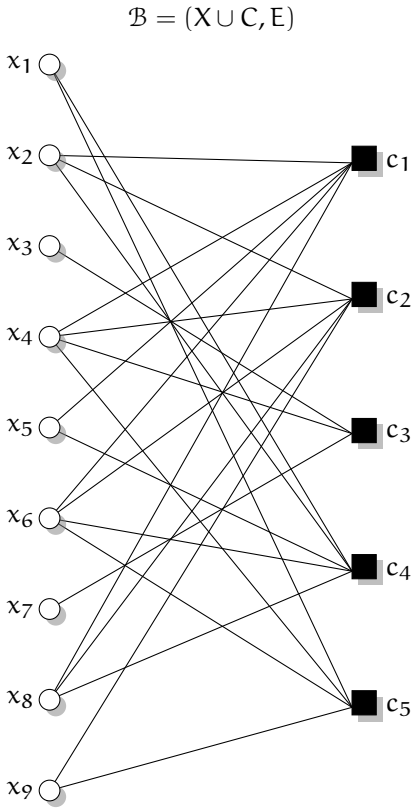


Figura 47: Grafo bipartito irregular asociado con un código LDPC irregular con matriz de chequeo de paridad H.

Además, contruimos el grafo bipartito $\mathcal{B} = (X \cup C, E)$ *correspondiente a la matriz* H*, que es irregular y donde* $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$ *es el conjunto con* 9 *nodos variable y* $C = \{c_1, c_2, c_3, c_4, c_5\}$ *es el conjunto con* 5 *nodos restricción, el cual ilustramos en la Figura 47. Al observar el grafo bipartito* $\mathcal{B}$ *vemos que los nodos adyacentes a cada nodo restricción son justamente los nodos variable asociados con las ecuaciones descritas anteriormente.*

Con ayuda del software ***Magma V2.14-15*** *[CBFS08] analicemos la matriz* H*.*

```
> M := KMatrixSpace(FiniteField(2),5,9);
> M1 := M ! [ 0,1,0,1,1,1,0,1,0,
              0,1,0,1,0,1,0,1,1,
              0,0,1,1,0,0,1,0,0,
              1,1,0,0,1,1,0,1,0,
              1,0,0,1,0,1,0,0,1 ];
> H := SparseMatrix(M1);
> H;

Sparse matrix with 5 rows and 9 columns over GF(2)

> C := LDPCCode(H);
> C;

[9, 4, 2] Linear Code over GF(2)
Generator matrix:
[1 0 0 1 0 0 1 1 0]
[0 1 0 0 0 0 0 1 0]
[0 0 1 0 0 0 1 0 0]
[0 0 0 0 1 1 0 0 1]

> CodeWordsC := [x : x in C];
> CodeWordsC;

[
    (0 0 0 0 0 0 0 0 0),
    (1 0 0 1 0 0 1 1 0),
    (1 1 0 1 0 0 1 0 0),
    (0 1 0 0 0 0 0 1 0),
    (0 1 1 0 0 0 1 1 0),
    (1 1 1 1 0 0 0 0 0),
    (1 0 1 1 0 0 0 1 0),
    (0 0 1 0 0 0 1 0 0),
    (0 0 1 0 1 1 1 0 1),
    (1 0 1 1 1 1 0 1 1),
    (1 1 1 1 1 1 0 0 1),
    (0 1 1 0 1 1 1 1 1),
    (0 1 0 0 1 1 0 1 1),
    (1 1 0 1 1 1 1 0 1),
    (1 0 0 1 1 1 1 1 1),
    (0 0 0 0 1 1 0 0 1)
]

> #C;

16

> IsLDPC(C);

true

> IsRegularLDPC(C);
```

```
false 0 0

> TannerGraph(C);

Graph
Vertex  Neighbours

1       13 14 ;
2       10 11 13 ;
3       12 ;
4       10 11 12 14 ;
5       10 13 ;
6       10 11 13 14 ;
7       12 ;
8       10 11 13 ;
9       11 14 ;
10      2 4 5 6 8 ;
11      2 4 6 8 9 ;
12      3 4 7 ;
13      1 2 5 6 8 ;
14      1 4 6 9 ;
```

Entonces, una vez realizado el análisis con Magma, los resultados obtenidos son que el código es un código LDPC de dimensión $\dim \mathcal{C}(\mathcal{B}) = 4$*, con palabras-código* x *de longitud* 9*, con distancia mínima* $d = 2$*, y, con diferentes pesos en las columnas* $w_{x_1} = 2, w_{x_2} = 3, w_{x_3} = 1, w_{x_4} = 4, w_{x_5} = 2, w_{x_6} = 4, w_{x_7} = 1, w_{x_8} = 3$ *y* $w_{x_9} = 2$*, y diferentes pesos en los renglones* $w_{c_1} = 5, w_{c_2} = 5, w_{c_3} = 3, w_{c_4} = 5$ *y* $w_{c_5} = 4$*. Es decir, un* $[9, 4, 2]$*-código LDPC irregular con tasa* $R = \frac{4}{9}$*, cuya matriz generadora es,*

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

con cardinalidad de $|\mathcal{C}(\mathcal{B})| = 16$ *palabras-código* x *que están enlistadas anteriormente. Además, al ser un código LDPC irregular el grafo bipartito asociado es un grafo bipartito irregular y en la parte final del análisis se obtienen los vecinos de los* 14 *nodos en total del grafo bipartito* $\mathcal{B}$ *y que son justamente los que se observan en la Figura 47. Por lo tanto, el* $[9, 4, 2]$*-código LDPC con matriz de chequeo de paridad* H *(4.7) y con representación gráfica bipartita* $\mathcal{B}$ *47, es el conjunto*

$$\mathcal{C}(\mathcal{B}) = \left\{ (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) \mid x_{b(1,j)} \oplus x_{b(2,j)} \oplus \cdots \oplus x_{b(d_{c_j},j)} = 0, 1 \leqslant j \leqslant 5 \right\}.$$

4.3 DISTRIBUCIÓN DE GRADO

La irregularidad de un código LDPC es usualmente descrita a través de las distribuciones de los grados de sus nodos variable y restricción, veamos la demostración del siguiente lema.

Lema 4.1. *El grado promedio de los nodos variable es* $a_x = \frac{1}{\int_0^1 \lambda(x)\,dx}$ *y el grado promedio de los nodos restricción es* $a_c = \frac{1}{\int_0^1 \rho(x)\,dx}$.

Demostración. Sea $\mathcal{B} = (X \cup C, E)$ un grafo bipartito irregular que es representado por la secuencia de los grados de los nodos variable $(l_1, ..., l_{d_x})$ y por la secuencia de los grados de los nodos restricción $(r_1, ..., r_{d_c})$, donde cada l_i y r_i son las fracciones de los n nodos variable y los m nodos restricción que tienen grado i, respectivamente, además, algunos de los $l_{i's}$ y $r_{i's}$ pueden ser cero; y d_x es el grado

máximo de los nodos variable y d_c es el grado máximo de los nodos resctricción, y se satisface que la suma de cada una de las secuencias es igual a 1. Sea λ_i la probabilidad de elegir al azar una arista que es adyacente a algún nodo variable de grado i, y sea ρ_i la probabilidad de que aleatoriamente elijamos una arista que es adyacente a algún nodo restricción de grado i, en otras palabras, λ_i y ρ_i son las fracciones de aristas conectadas a los nodos variable y nodos restricción de grado i, respectivamente. Así,

$$\lambda_i = \frac{il_i}{\sum_{i=1}^{d_x} il_i} \quad i = 1, \ldots, d_x \qquad y \qquad \rho_i = \frac{ir_i}{\sum_{i=1}^{d_c} ir_i} \quad i = 1, \cdots, d_c,$$

podemos observar que,

$$\sum_i \lambda_i = \sum_{i=1}^{d_x} \frac{il_i}{\sum_{i=1}^{d_x} il_i} = \frac{\sum_{i=1}^{d_x} il_i}{\sum_{i=1}^{d_x} il_i} = 1$$

$$\sum_i \rho_i = \sum_{i=1}^{d_c} \frac{ir_i}{\sum_{i=1}^{d_c} ir_i} = \frac{\sum_{i=1}^{d_c} ir_i}{\sum_{i=1}^{d_c} ir_i} = 1$$

Ahora definamos las distribuciones de grado arista $\lambda(x)$ y $\rho(x)$ que son polinomios de la siguiente forma,

$$\lambda(x) = \sum_i \lambda_i x^{i-1} \qquad y \qquad \rho(x) = \sum_i \rho_i x^{i-1}.$$

Para obtener el grado promedio de los nodos variable, contamos el número de aristas de grado i de los nodos variable, los cuales deben ser igual a $\lambda_i n a_x$. Dado que el número de nodos variable de grado i del grafo es $\frac{\lambda_i n a_x}{i}$, obtenemos que l_i la fracción del nodo variable de grado i, está dada por $l_i = \frac{\lambda_i a_x}{i}$. Como se satisface que $\sum_i l_i = \sum_i \frac{\lambda_i a_x}{i} = 1$ resulta que $a_x = \frac{1}{\sum_i \frac{\lambda_i}{i}}$. Observemos que $\int_0^1 \lambda(x)dx = \int_0^1 \sum_{i=1}^{d_x} \lambda_i x^{i-1} dx = \sum_{i=1}^{d_x} \int_0^1 \lambda_i x^{i-1} dx = \sum_{i=1}^{d_x} \frac{\lambda_i}{i} x^i|_{x=0}^{x=1} = \sum_{i=1}^{d_x} \frac{\lambda_i}{i}$, por lo tanto, $a_x = \frac{1}{\int_0^1 \lambda(x)dx}$. De manera similar, para el grado promedio de los nodos restricción, contamos el número de aristas de grado i de los nodos restricción, estos deben ser igual a $\rho_i m a_c$. Luego, el número de nodos restricción de grado i del grafo es $\frac{\rho_i m a_c}{i}$, con lo que obtenemos que r_i la fracción del nodo restricción de grado i, está dada por $r_i = \frac{\rho_i a_c}{i}$. Dado que también se cumple que $\sum_i \rho_i = \sum_i \frac{\rho_i a_c}{i} = 1$ entonces $a_c = \frac{1}{\sum_i \frac{\rho_i}{i}}$. Nótese que $\int_0^1 \rho(x)dx = \int_0^1 \sum_{i=1}^{d_c} \rho_i x^{i-1} dx = \sum_{i=1}^{d_c} \int_0^1 \rho_i x^{i-1} dx = \sum_{i=1}^{d_c} \frac{\rho_i}{i} x^i|_{x=0}^{x=1} = \sum_{i=1}^{d_c} \frac{\rho_i}{i}$, por lo tanto, $a_c = \frac{1}{\int_0^1 \rho(x)dx}$. □

Ejemplo 4.3. *Consideremos el Ejemplo 4.2 donde trabajamos un $[9, 4, 2]$-código LDPC irregular dado por su matriz de chequeo de paridad H (4.7) y con su respectivo grafo bipartito $\mathcal{B} = (X \cup C, E)$ irregular con 9 nodos variable, 5 nodos restricción y 22 aristas, ilustrado en la Figura 47. Vamos a calcular y analizar los diferentes conceptos indicados en la demostración del Lema 4.1.*

En las Tablas 13 y 14 damos los grados de cada nodo del grafo bipartito $\mathcal{B}$, así como el grado máximo y promedio,

Nodo variable x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
Grado d_{x_i}	2	3	1	4	2	4	1	3	2
Grado máximo d_x	4								
Grado promedio a_x	$\frac{22}{9}$								

Tabla 13: Grado de cada nodo variable.

Nodo restricción c_i	c_1	c_2	c_3	c_4	c_5
Grado d_{c_i}	5	5	3	5	4
Grado máximo d_c	5				
Grado promedio a_c	$\frac{22}{5}$				

Tabla 14: Grado de cada nodo restricción.

Calculemos los $l_{i's}$ y $r_{i's}$ que son las fracciones de los nodos variable y restricción de grado i,

Secuencias de los grados	
Nodos variable	$(l_1, l_2, l_3, l_4) = (\frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{2}{9})$
Nodos restricción	$(r_1, r_2, r_3, r_4, r_5) = (0, 0, \frac{1}{5}, \frac{1}{5}, \frac{3}{5})$

Tabla 15: Fracciones de los nodos variable y los nodos restricción.

De la Tabla 15, consideremos la primera fracción $l_1 = \frac{2}{9}$, en sí, lo que está fracción indica es que hay 2 nodos variable que tienen grado 1 de los 9 nodos variable del grafo bipartito $\mathcal{B}$. Para la fracción $l_2 = \frac{3}{9}$, se tiene que hay 3 nodos variable de grado 2 de los 9 nodos variable, y así sucesivamente con las fracciones restantes, podemos verificar lo anteriormente dicho observando el grafo bipartito $\mathcal{B}$ 47.
De la misma Tabla 15, tomemos $r_1 = 0$, este valor indica que hay 0 nodos restricción que tienen grado 1 de los 5 nodos restricción en el grafo bipartito. Para $r_3 = \frac{1}{5}$ decimos que hay 1 nodo restricción de grado 3 de los 5 nodos restricción, todo esto se puede verificar revisando el grafo bipartito $\mathcal{B}$ 47.

Ahora calculemos las siguientes sumas,

$$\sum_{i=1}^{4} l_i = l_1 + l_2 + l_3 + l_4 = \frac{2}{9} + \frac{3}{9} + \frac{2}{9} + \frac{2}{9} = 1$$

$$\sum_{i=1}^{5} r_i = r_1 + r_2 + r_3 + r_4 + r_5 = 0 + 0 + \frac{1}{5} + \frac{1}{5} + \frac{3}{5} = 1$$

$$\sum_{i=1}^{4} i l_i = 1l_1 + 2l_2 + 3l_3 + 4l_4 = 1 \cdot \frac{2}{9} + 2 \cdot \frac{3}{9} + 3 \cdot \frac{2}{9} + 4 \cdot \frac{2}{9} = \frac{2}{9} + \frac{6}{9} + \frac{6}{9} + \frac{8}{9} = \frac{22}{9}$$

$$\sum_{i=1}^{5} i r_i = 1r_1 + 2r_2 + 3r_3 + 4r_4 + 5r_5 = 1 \cdot 0 + 2 \cdot 0 + 3 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} + 5 \cdot \frac{3}{5} = 0 + 0 + \frac{3}{5} + \frac{4}{5} + \frac{15}{5} = \frac{22}{5}$$

Calculemos λ_i y ρ_i las fracciones de aristas conectadas a los nodos variable y nodos restricción de grado i,

$$\lambda_1 = \frac{1l_1}{\Sigma_{i=1}^{4} i l_i} = \frac{\frac{2}{9}}{\frac{22}{9}} = \frac{2}{22}$$

$$\lambda_2 = \frac{2l_2}{\Sigma_{i=1}^{4} i l_i} = \frac{\frac{6}{9}}{\frac{22}{9}} = \frac{6}{22}$$

$$\lambda_3 = \frac{3l_3}{\Sigma_{i=1}^{4} i l_i} = \frac{\frac{6}{9}}{\frac{22}{9}} = \frac{6}{22}$$

$$\lambda_4 = \frac{4l_4}{\Sigma_{i=1}^{4} i l_i} = \frac{\frac{8}{9}}{\frac{22}{9}} = \frac{8}{22}$$

De los calculos anteriores, tomemos $\lambda_1 = \frac{2}{22}$, esta fracción indica que hay 2 aristas incidentes a los nodos variable con grado 1 de las 22 aristas totales en el grafo bipartito. Para $\lambda_4 = \frac{8}{22}$ decimos que hay 8 aristas incidentes a los nodos variable de grado 4 de las 22 aristas. Se puede interpretar de forma similar los valores restantes.

$$\rho_1 = \frac{1r_1}{\Sigma_{i=1}^{5} ir_i} = \frac{0}{\frac{22}{5}} = 0$$

$$\rho_2 = \frac{2r_2}{\Sigma_{i=1}^{5} ir_i} = \frac{0}{\frac{22}{5}} = 0$$

$$\rho_3 = \frac{3r_3}{\Sigma_{i=1}^{5} ir_i} = \frac{\frac{3}{5}}{\frac{22}{5}} = \frac{3}{22}$$

$$\rho_4 = \frac{4r_4}{\Sigma_{i=1}^{5} ir_i} = \frac{\frac{4}{5}}{\frac{22}{5}} = \frac{4}{22}$$

$$\rho_5 = \frac{5r_5}{\Sigma_{i=1}^{5} ir_i} = \frac{\frac{15}{5}}{\frac{22}{5}} = \frac{15}{22}$$

Ahora, tomemos el valor de $\rho_2 = 0$, esto indica que no hay aristas incidentes a algún nodo restricción con grado 2, ni siquiera hay nodos restricción de grado 2. Sea $\rho_3 = \frac{3}{22}$, podemos decir que, hay 3 arista incidente a los nodos restricción de grado 3 de las 22 aristas. De igual manera se pueden interpretar los demás valores.

Veamos que se cumple que la suma de todos los $\lambda_{i's}$, así, como la suma de todos los $\rho_{i's}$, es igual a 1.

$$\sum_{i=1}^{4} \lambda_i = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = \frac{2}{22} + \frac{6}{22} + \frac{6}{22} + \frac{8}{22} = 1$$

$$\sum_{i=1}^{5} \rho_i = \rho_1 + \rho_2 + \rho_3 + \rho_4 + \rho_5 = 0 + 0 + \frac{3}{22} + \frac{4}{22} + \frac{15}{22} = 1$$

Ahora calculemos los polinomios $\lambda(x)$ y $\rho(x)$ que son las distribuciones de grado arista,

$$\begin{aligned}
\lambda(x) &= \sum_{i=1}^{4} \lambda_i x^{i-1} \\
&= \lambda_1 x^0 + \lambda_2 x^1 + \lambda_3 x^2 + \lambda_4 x^3 \\
&= \frac{2}{22} + \frac{6}{22}x + \frac{6}{22}x^2 + \frac{8}{22}x^3 \\
&= \frac{1}{11} + \frac{3}{11}x + \frac{3}{11}x^2 + \frac{4}{11}x^3
\end{aligned}$$

$$\begin{aligned}
\rho(x) &= \sum_{i=1}^{5} \rho_i x^{i-1} \\
&= \rho_1 x^0 + \rho_2 x^1 + \rho_3 x^2 + \rho_4 x^3 + \rho_5 x^4 \\
&= 0 + 0 + \frac{3}{22}x^2 + \frac{4}{22}x^3 + \frac{15}{22}x^4 \\
&= \frac{3}{22}x^2 + \frac{4}{22}x^3 + \frac{15}{22}x^4
\end{aligned}$$

Finalmente, de acuerdo al resultado del Lema 4.1, tenemos que los grados promedios pueden ser calculados de la siguiente manera,

$$a_x = \frac{1}{\int_0^1 \lambda(x)dx} \qquad y \qquad a_c = \frac{1}{\int_0^1 \rho(x)dx}, \tag{4.13}$$

entonces,

$$\begin{aligned}
\int_0^1 \lambda(x)dx &= \int_0^1 (\frac{1}{11} + \frac{3}{11}x + \frac{3}{11}x^2 + \frac{4}{11}x^3)dx \\
&= \frac{1}{11}x + \frac{3}{22}x^2 + \frac{3}{33}x^3 + \frac{4}{44}x^4\Big|_0^1 \\
&= \frac{1}{11}\cdot 1 + \frac{3}{22}\cdot 1^2 + \frac{1}{11}\cdot 1^3 + \frac{1}{11}\cdot 1^4 - (\frac{1}{11}\cdot 0 + \frac{3}{22}\cdot 0^2 + \frac{1}{11}\cdot 0^3 + \frac{1}{11}\cdot 0^4) \\
&= \frac{1}{11} + \frac{3}{22} + \frac{1}{11} + \frac{1}{11} - 0 \\
&= \frac{9}{22}
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \rho(x)dx &= \int_0^1 (\frac{3}{22}x^2 + \frac{4}{22}x^3 + \frac{15}{22}x^4)dx \\
&= \frac{3}{66}x^3 + \frac{4}{88}x^4 + \frac{15}{110}x^5\Big|_0^1 \\
&= \frac{1}{22}\cdot 1^3 + \frac{1}{22}\cdot 1^4 + \frac{3}{22}\cdot 1^5 - (\frac{1}{22}\cdot 0^3 + \frac{1}{22}\cdot 0^4 + \frac{3}{22}\cdot 0^5) \\
&= \frac{1}{22} + \frac{1}{22} + \frac{3}{22} - 0 \\
&= \frac{5}{22}
\end{aligned}$$

luego,

$$\frac{1}{\int_0^1 \lambda(x)dx} = \frac{1}{\frac{9}{22}} = \frac{22}{9} = a_x \qquad \textit{(ver la Tabla 13)}$$

$$\frac{1}{\int_0^1 \rho(x)dx} = \frac{1}{\frac{5}{22}} = \frac{22}{5} = a_c \qquad \textit{(ver la Tabla 14)}$$

Vemos que se satisfacen los resultados del Lema 4.13, ya que los valores coinciden con el grado promedio de los nodos variable y restricción, respectivamente.

4.4 COTAS PARA LA DISTANCIA MÍNIMA DE UN CÓDIGO LINEAL POR ANÁLISIS DEL GRAFO BIPARTITO

Ahora analizamos dos cotas inferiores simples para la distancia mínima d, de un código LDPC, $\mathcal{C}(\mathcal{B})$, (d_x, d_c)-regular con matriz de chequeo de paridad H y su respectivo grafo bipartito $\mathcal{B}$, tales teoremas fueron originalmente probados por Tanner [Tano1].

4.4.1 Primera cota por análisis de los nodos variable

La primera cota examina la relación entre los nodos variable en una palabra-código x de peso mínimo $w_x = d$ y la matriz de chequeo de paridad H. La cota es expresada en términos de los valores propios de la matriz $H^T H$ pues estos valores propios están relacionados con los valores propios de la matriz de adyacencia A del grafo bipartito, ver Lemas 3.3 y 3.5. Además, como $H^T H$ es una matriz simétrica y de valores reales entonces sus valores propios son números reales, ver Teorema 3.1, y consideramos a μ_1 como el valor propio más grande de todos, el cual cumple con muchas propiedades que se revisaron en la sección 3.3 del capítulo 3.

Teorema 4.1. *Si* $\mathcal{B} = (X \cup C, E)$ *es un grafo bipartito conexo* (d_x, d_c)*-regular con* n *nodos variable de grado* d_x *y* m *nodos restricción de grado* d_c*, la distancia mínima* d *del código* $\mathcal{C}(\mathcal{B})$ *satisface que*

$$d \geqslant \frac{n(2d_x - \mu_2)}{(\mu_1 - \mu_2)} \tag{4.14}$$

donde μ_1 *es el valor propio más grande y* μ_2 *es el segundo valor propio más grande de la matriz* $H^T H$ *con* H *la matriz de chequeo de paridad del código* $\mathcal{C}(\mathcal{B})$*.*

Demostración. Sean μ_i $1 \leqslant i \leqslant s$ los distintos valores propios reales de la matriz $H^T H$ donde H es la matriz de chequeo de paridad del código, y tales que los valores propios son ordenados $\mu_1 > \mu_2 > \cdots > \mu_s$. Como $\mathcal{B}$ es un grafo bipartito regular, $v_1 = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)^T$ es el vector propio de longitud n con su correspondiente valor propio $\mu_1 = d_x d_c$ de multiplicidad $m_1 = 1$ de la matriz $H^T H$, ver Proposiciones 3.2 y 3.3. Dado $x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{C}(\mathcal{B})$ un vector de valores reales de longitud n correspondiente a una palabra-código de peso mínimo d, recordar que para un código lineal el peso mínimo coincide con la distancia mínima, Teorema 2.2. Y sea y_i la proyección de x sobre el i-ésimo espacio propio.

Claramente, como x es una palabra-código de peso mínimo con entradas con valor 0 y exactamente d entradas con valor 1 se cumple que

$$x^T \cdot x = \|x\|^2 = d. \tag{4.15}$$

Ahora, de acuerdo al proceso de Gram-Schmidt, Teorema 3.11, y la normalización tenemos el vector ortonormal $u_1 = v_1 = \frac{1}{\sqrt{n}}(1, \ldots, 1)^T$. Como $H^T H$ es una matriz simétrica entonces es diagonalizable ortonormalmente, ver Teorema 3.12. Y se satisface que la matriz de proyección P_1 sobre el primer espacio propio es de la forma

$$P_1 = QD_1Q^T \qquad \text{(ver Observación 3.3)}$$

donde Q es la matriz formada por una base ortonormal, y D_1 es la matriz asociada al primer valor propio $\mu_1 = d_x d_c$ de multiplicidad $m_1 = 1$ que en cuya diagonal tiene en la primer componente una matriz identidad de tamaño 1×1 y el resto de sus entradas es 0, es decir,

$$Q = \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix}, \quad Q^T = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix}, \quad D_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Luego,

$$\begin{aligned} P_1 &= QD_1Q^T \\ &= \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \\ &= \begin{pmatrix} u_{11} & 0 & \cdots & 0 \\ u_{12} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \\ &= \begin{pmatrix} u_{11}^2 & u_{11}u_{12} & \cdots & u_{11}u_{1n} \\ u_{12}u_{11} & u_{12}^2 & \cdots & u_{12}u_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n}u_{11} & u_{1n}u_{12} & \cdots & u_{1n}^2 \end{pmatrix}. \end{aligned}$$

Entonces, como el vector ortonormal es $u_1 = \frac{1}{\sqrt{n}}(1,1,\ldots,1)^T$, se sigue que la matriz de proyección P_1 es

$$\begin{aligned} P1 &= \begin{pmatrix} u_{11}^2 & u_{11}u_{12} & \cdots & u_{11}u_{1n} \\ u_{12}u_{11} & u_{12}^2 & \cdots & u_{12}u_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n}u_{11} & u_{1n}u_{12} & \cdots & u_{1n}^2 \end{pmatrix} \\ &= \begin{pmatrix} (\frac{1}{\sqrt{n}})^2 & \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} & (\frac{1}{\sqrt{n}})^2 & \cdots & \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} & \cdots & (\frac{1}{\sqrt{n}})^2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \end{aligned}$$

Después al obtener y_1 la proyección de x sobre el primer espacio propio tenemos que

$$y_1 = P_1 x = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{n}\begin{pmatrix} d \\ d \\ \vdots \\ d \end{pmatrix} = \frac{d}{n}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

pues x es una palabra-código de peso mínimo d. Así, la proyección y_1 de x es $y_1 = \frac{d}{n}(1,1,\ldots,1)^T$, luego,

$$\left|\left|y_1\right|\right|^2 = \left|\left|\frac{d}{n}(1,1,\ldots,1)^T\right|\right|^2 = \left|\frac{d}{n}\right|^2\left|\left|(1,1,\ldots,1)^T\right|\right|^2 = \frac{d^2}{n^2}\sum_{i=1}^{n} 1^2 = \frac{d^2}{n^2}\cdot n = \frac{d^2}{n}. \tag{4.16}$$

Hx asigna una distribución de peso para los nodos chequeo en $\mathcal{B}$, es decir, sea ω_i el peso sobre el i-ésimo renglón paridad definido por Hx que representa el número de nodos variable activos adyacentes al i-ésimo nodo restricción en el grafo bipartito $\mathcal{B}$ ("activo"se refiere a los nodos variable distintos de cero en el correspondiente vector x), esto es,

$$\begin{aligned} Hx &= \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^{n} h_{1i}x_i \\ \sum_{i=1}^{n} h_{2i}x_i \\ \vdots \\ \sum_{i=1}^{n} h_{mi}x_i \end{pmatrix} \\ &= \begin{pmatrix} h_1 \cdot x \\ h_2 \cdot x \\ \vdots \\ h_m \cdot x \end{pmatrix} \qquad (\text{donde} \quad h_i = (h_{i1}, h_{i2}, \ldots, h_{in}) \quad 1 \leqslant i \leqslant m) \\ &= \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_m \end{pmatrix}. \end{aligned}$$

Entonces $\|Hx\|^2 = \sum_{i=1}^{r} \omega_i^2$. Además, $\sum_{i=1}^{m} \omega_i = \sum_{i=1}^{m} h_i \cdot x = h_1 \cdot x + h_2 \cdot x + \cdots + h_m \cdot x = (h_1 + h_2 + \cdots + h_m) \cdot x$. Ahora, como H es la matriz de chequeo de paridad de un código regular, tenemos que cada renglón de H tiene peso $w_c = d_c$ y cada columna de H tiene peso $w_x = d_x$, esto es, los pesos de la matriz H coinciden con los grados del grafo bipartito $\mathcal{B}$. Luego, al sumar la primera entrada de cada h_i, estaríamos sumando las componente de la primer columna en H, así, la suma resulta ser igual al peso de la columna la cual es $w_x = d_x$, lo mismo resultaría para las demás entradas. Por lo que se sigue $(h_1 + h_2 + \cdots + h_m) \cdot x = (d_x, d_x, \ldots, d_x) \cdot x = d_x(1, 1, \ldots, 1) \cdot x = d_x d$, por lo tanto, $\sum_{i=1}^{m} \omega_i = d d_x$. Nótese que no todos los pesos $\omega_{i's}$ son cero y cada $\omega_i \neq 0$ debería tener valor de al menos 2, porque al menos dos nodos variable deben ser adyacentes a algún nodo restricción para asegurar que tenemos una palabra-código, es decir, cada una de las palabras-código excepto la palabra-código cero tienen al menos dos nodos variables activos pues si esto no ocurre no se tendría una palabra-código. Como x es una palabra-código distinta de cero, y además $0 \leqslant \omega_i \leqslant d_c$, por consiguiente se satisface que

$$\|Hx\|^2 = \sum_{i=1}^{m} \omega_i^2 \geqslant 2 \sum_{i=1}^{m} \omega_i = 2 d d_x. \tag{4.17}$$

Por otra parte, usando las propiedades que se satisfacen del Teorema de descomposición espectral 3.3, tenemos que

$$\begin{aligned}
\|Hx\|^2 &= < Hx, Hx > \\
&= (Hx)^T (Hx) \\
&= x^T H^T H x \\
&= x^T (\mu_1 P_1 + \mu_2 P_2 + \cdots + \mu_s P_s) x \\
&= x^T (\mu_1 P_1) x + x^T (\mu_2 P_2) x + \cdots + x^T (\mu_s P_s) x \\
&= \mu_1 (x^T P_1 x) + \mu_2 (x^T P_2 x) + \cdots + \mu_s (x^T P_s x) \\
&= \mu_1 (x^T P_1 P_1 x) + \mu_2 (x^T P_2 P_2 x) + \cdots + \mu_s (x^T P_s P_s x) \\
&= \mu_1 (x^T P_1^T P_1 x) + \mu_2 (x^T P_2^T P_2 x) + \cdots + \mu_s (x^T P_s^T P_s x) \\
&= \mu_1 (P_1 x)^T (P_1 x) + \mu_2 (P_2 x)^T (P_2 x) + \cdots + \mu_s (P_s x)^T (P_s x) \\
&= \mu_1 < P_1 x, P_1 x > + \mu_2 < P_2 x, P_2 x > + \cdots + \mu_s < P_s x, P_s x > \\
&= \mu_1 \|P_1 x\|^2 + \mu_2 \|P_2 x\|^2 + \cdots + \mu_s \|P_s x\|^2 \\
&= \mu_1 \|y_1\|^2 + \mu_2 \|y_2\|^2 + \cdots + \mu_s \|y_s\|^2 \\
&= \sum_{i=1}^{s} \mu_i \|y_i\|^2
\end{aligned}$$

Entonces, $\|Hx\|^2 = \sum_{i=1}^{s} \mu_i \|y_i\|^2$. Además,

$$\begin{aligned}
\sum_{i=1}^{s} \mu_i \|y_i\|^2 &= \mu_1 \|y_1\|^2 + \sum_{i=2}^{s} \mu_i \|y_i\|^2 \\
&= \mu_1 \frac{d^2}{n} + \sum_{i=2}^{s} \mu_i \|y_i\|^2 \qquad \text{(por (4.16))}
\end{aligned} \tag{4.18}$$

Observemos que se cumple lo siguiente usando las propiedades del Teorema espectral 3.3,

$$\begin{aligned}
\sum_{i=1}^{s} \|y_i\|^2 &= \|y_1\|^2 + \|y_2\|^2 + \cdots + \|y_s\|^2 \\
&= \|P_1 x\|^2 + \|P_2 x\|^2 + \cdots + \|P_s x\|^2 \\
&= < P_1 x, P_1 x > + < P_2 x, P_2 x > + \cdots + < P_s x, P_s x > \\
&= (P_1 x)^T (P_1 x) + (P_2 x)^T (P_2 x) + \cdots + (P_s x)^T (P_s x) \\
&= x^T P_1^T P_1 x + x^T P_2^T P_2 x + \cdots + x^T P_s^T P_s x \\
&= x^T P_1 P_1 x + x^T P_2 P_2 x + \cdots + x^T P_s P_s x \\
&= x^T P_1^2 x + x^T P_2^2 x + \cdots + x^T P_s^2 x
\end{aligned}$$

$$
\begin{aligned}
&= x^T P_1 x + x^T P_2 x + \cdots + x^T P_s x \\
&= x^T (P_1 + P_2 + \cdots + P_s) x \\
&= x^T I x \\
&= x^T x \\
&= < x, x > \\
&= \|x\|^2
\end{aligned}
$$

Así, $\sum_{i=1}^{s} \|y_i\|^2 = \|x\|^2$, luego $\|y_1\|^2 + \sum_{i=2}^{s} \|y_i\|^2 = \|x\|^2$, entonces $\sum_{i=2}^{s} \|y_i\|^2 = \|x\|^2 - \|y_1\|^2$. Regresando a la ecuación (4.18),

$$
\begin{aligned}
\sum_{i=1}^{s} \mu_i \|y_i\|^2 &= \mu_1 \frac{d^2}{n} + \sum_{i=2}^{s} \mu_i \|y_i\|^2 \\
&\leqslant \mu_1 \frac{d^2}{n} + \mu_2 \sum_{i=2}^{s} \|y_i\|^2 \qquad (\text{pues } \mu_2 > \mu_3 > \cdots > \mu_s) \\
&= \mu_1 \frac{d^2}{n} + \mu_2 (\|x\|^2 - \|y_1\|^2) \\
&= \mu_1 \frac{d^2}{n} + \mu_2 (d - \frac{d^2}{n}) \qquad (\text{por (4.15) y (4.16)})
\end{aligned}
$$

Por lo tanto,

$$\|Hx\|^2 \leqslant \mu_1 \frac{d^2}{n} + \mu_2 (d - \frac{d^2}{n}) \tag{4.19}$$

Combinando las desigualdades (4.17) y (4.19) tenemos lo siguiente,

$$2dd_x \leqslant \|Hx\|^2 \leqslant \mu_1 \frac{d^2}{n} + \mu_2 (d - \frac{d^2}{n})$$

entonces por transitividad,

$$
\begin{aligned}
2dd_x \leqslant \mu_1 \frac{d^2}{n} + \mu_2 (d - \frac{d^2}{n}) \quad &\Rightarrow \quad 2d_x \leqslant \mu_1 \frac{d}{n} + \mu_2 (1 - \frac{d}{n}) \\
&\Rightarrow \quad 2d_x n \leqslant \mu_1 d + \mu_2 (n - d) \\
&\Rightarrow \quad 2d_x n - \mu_2 n \leqslant (\mu_1 - \mu_2) d \\
&\Rightarrow \quad \frac{(2d_x - \mu_2) n}{(\mu_1 - \mu_2)} \leqslant d
\end{aligned}
$$

Por lo tanto, la distancia mínima está acotada por $d \geqslant \frac{n(2d_x - \mu_2)}{(\mu_1 - \mu_2)}$. □

Veamos el siguiente ejemplo.

Ejemplo 4.4. *Sea $\mathcal{C}(\mathcal{B})$ el $[8,5]$-código LDPC $(2,4)$-regular con matriz de chequeo de paridad H (4.2) y grafo bipartito $\mathcal{B} = (X \cup C, E)$, Figura 46, con 8 nodos variable y 4 nodos restricción del Ejemplo 4.1. El objetivo es acotar la distancia mínima d del $[8,5]$-código LDPC $\mathcal{C}(\mathcal{B})$. De acuerdo al Teorema 4.1 para acotar la distancia mínima d del código LDPC $\mathcal{C}(\mathcal{B})$ necesitamos conocer el número de nodos variable n, el grado de los nodos variable d_x, y el primer y segundo valor propio más grandes μ_1, μ_2 de la matriz $H^T H$ que están asociados con los valores propios del grafo bipartito $\mathcal{B}$.*

Sabemos por el Ejemplo 4.1 que para dicho código LDPC $n = 8$, $d_x = 2$, y por el Ejemplo 3.1 obtenemos en la salida (%07) que los valores propios de la matriz $H^T H$ son $\mu_1 = 8$ de multiplicidad $m_1 = 1$, $\mu_2 = 4$ de multiplicidad $m_2 = 1$, $\mu_3 = 2$ con multiplicidad $m_3 = 2$ y $\mu_4 = 0$ de multiplicidad $m_4 = 4$, ordenados de tal forma que $\mu_1 > \mu_2 > \mu_3 > \mu_4$. Así,

$$d \geqslant \frac{n(2d_x - \mu_2)}{\mu_1 - \mu_2} = \frac{8(2 \cdot 2 - 4)}{8 - 4} = \frac{8(0)}{4} = 0,$$

entonces la distancia mínima es mayor o igual a cero $d \geqslant 0$, esta cota no nos proporciona mucha información, aunque sabemos que esto si se satisface, pues la distancia mínima es $d = 2$ para el $[8,5]$-código LDPC.

4.4.2 Segunda cota por análisis de los nodos restricción

La segunda cota considera la relación entre los nodos restricción adyacentes a los nodos variable de una palabra-código x de peso mínimo $w_x = d$ y la matriz H^T. La cota es expresada en términos de los valores propios de la matriz HH^T, podemos notar que, los valores propios no cero de la matriz H^TH y HH^T son los mismos, ver Lema 3.3.

Teorema 4.2. *Si $\mathcal{B} = (X \cup C, E)$ es un grafo bipartito conexo (d_x, d_c)-regular con n nodos variable de grado d_x, y m nodos restricción de grafo d_c, la distancia mínima d del código $\mathcal{C}(\mathcal{B})$ satisface que*

$$d \geqslant \frac{2n(2d_x + d_c - 2 - \mu_2)}{d_c(\mu_1 - \mu_2)} \tag{4.20}$$

donde μ_1 es el valor propio más grande y μ_2 es el segundo valor propio más grande de la matriz HH^T con H la matriz de chequeo de paridad del código $\mathcal{C}(\mathcal{B})$.

Demostración. Como los valores propios distintos de cero de H^TH y HH^T son los mismos, entonces μ_i, $1 \leqslant i \leqslant r$, son los valores propios de HH^T tales que $\mu_1 > \mu_2 > \cdots > \mu_r$.
Sea el primer vector propio $v_1 = \frac{1}{\sqrt{m}}(1, 1, \ldots, 1)^T$ de longitud m correspondiente al valor propio $\mu_1 = d_x d_c$ de multiplicidad $m_1 = 1$ de la matriz HH^T, esto se debe a que tenemos un grafo bipartito regular $\mathcal{B}$, ver Proposiciones 3.2 y 3.3. Definamos a $c = (c_1, c_2, \ldots, c_m)^T$ como un vector de valores reales de longitud m de peso w, donde tiene valor 1 en la entrada donde el nodo restricción es adyacente a cualquier nodo variable de alguna palabra-código x no cero con peso mínimo d y tiene el valor de 0 en otro caso. Y sea z_i la proyección de c sobre el i-ésimo espacio propio.
Como w es el número de $1's$ en c, entonces

$$c^T \cdot c = \|c\|^2 = w. \tag{4.21}$$

Ahora, usando el proceso de Gram-Schmidt, Teorema 3.11, y la normalización el primer vector ortonormal $u_1 = v_1 = \frac{1}{\sqrt{m}}(1, \ldots, 1)^T$. Como HH^T es una matriz simétrica entonces es diagonalizable ortonormalmente, ver Teorema 3.12. Y se satisface que la matriz de proyección P_1 sobre el primer espacio propio es de la forma

$$P_1 = QD_1Q^T \qquad \text{(ver Observación 3.3)}$$

donde Q es la matriz formada por una base ortonormal, y D_1 es la matriz asociada al primer valor propio $\mu_1 = d_x d_c$ de multiplicidad $m_1 = 1$ que en cuya diagonal la primer componente es una matriz identidad de tamaño 1×1 y el resto de sus entradas es 0. Luego, la matriz de proyección P_1 es

$$P_1 = \begin{pmatrix} u_{11}^2 & u_{11}u_{12} & \cdots & u_{11}u_{1m} \\ u_{12}u_{11} & u_{12}^2 & \cdots & u_{12}u_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1m}u_{11} & u_{1m}u_{12} & \cdots & u_{1m}^2 \end{pmatrix}$$

entonces, como el primer vector ortonormal es $u_1 = \frac{1}{\sqrt{m}}(1, 1, \ldots, 1)^T$, se sigue que,

$$P_1 = \frac{1}{m}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

luego, z_1 la proyección de c sobre el primer espacio propio es

$$z_1 = P_1 c = \frac{1}{m}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \frac{w}{m}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

pues c es un vector de peso w. Así, la proyección z_1 de c es $z_1 = \frac{w}{m}(1,1\ldots,1)^T$, luego,

$$||z_1||^2 = \left|\left|\frac{w}{m}(1,1\ldots,1)^T\right|\right|^2 = \left|\frac{w}{m}\right|^2\left|\left|(1,1\ldots,1)^T\right|\right|^2 = \frac{w^2}{m^2}\sum_{i_1}^{m} 1^2 = \frac{w^2}{m^2}\cdot m = \frac{w^2}{m} \tag{4.22}$$

$H^T c$ asigna una distribución de peso para los nodos variable en $\mathcal{B}$, es decir, sea ν_i el peso sobre el i-ésimo renglón bit definido por $H^T c$ que representa el número de nodos restricción activos adyacentes al i-ésimo nodo variable en el grafo bipartito $\mathcal{B}$ ("activo"se refiere a los nodos restricción distintos de cero en el correspondiente vector c), esto es,

$$\begin{aligned}
H^T c &= \begin{pmatrix} h_{11} & h_{21} & \cdots & h_{m1} \\ h_{12} & h_{22} & \cdots & h_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1n} & h_{2n} & \cdots & h_{mn} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^{m} h_{i1}c_i \\ \sum_{i=1}^{m} h_{i2}c_i \\ \vdots \\ \sum_{i=1}^{m} h_{in}c_i \end{pmatrix} \\
&= \begin{pmatrix} h_1 \cdot c \\ h_2 \cdot c \\ \vdots \\ h_n \cdot c \end{pmatrix} \qquad (h_i = (h_{1i}, h_{2i}, \ldots, h_{mi}) \quad 1 \leqslant i \leqslant n) \\
&= \begin{pmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_n \end{pmatrix}.
\end{aligned}$$

Entonces, $\|H^T c\|^2 = \sum_{i=1}^{n} \nu_i^2$. Notar que no todos los $\nu_{i's}$ son cero, y además, $0 \leqslant \nu_i \leqslant d_x$. Ahora, sea $u_j(l)$ el número de nodos variable con peso l en $H^T c$ que son adyacentes al j-ésimo nodo restricción activo, para $1 \leqslant l \leqslant d_x$. Cada nodo restricción activo es adyacente a al menos 2 nodos variable, por lo tanto $u_j(d_x) \geqslant 2$.

La suma de los pesos al cuadrado puede ser calculada al sumar las aristas incidentes a los nodos variable adyacentes a todos los nodos de restricción activos. La siguiente sumatoria indica la suma de las aristas incidentes a los nodos variable adyacentes al j-ésimo nodo restricción,

$$\begin{aligned}
\sum_{l=1}^{d_x} u_j(l)l &= u_j(d_x)d_x + \sum_{l=1}^{d_x-1} u_j(l)l \\
&\geqslant 2d_x + d_c - 2,
\end{aligned}$$

ya que la segunda sumatoria es acotada inferiormente por $d_c - 2$. Luego, hay w nodos de paridad activos en el vector c, por lo tanto,

$$\|H^T c\|^2 = \sum_{i=1}^{n} \nu_i^2 \geqslant w(2d_x + d_c - 2). \tag{4.23}$$

Por otro lado, usando las propiedades que se satisfacen del Teorema de descomposición espectral 3.3, tenemos que

$$\begin{aligned}
\|H^T c\|^2 &= < H^T c, H^T c > \\
&= (H^T c)^T (H^T c) \\
&= c^T (H^T)^T H^T c \\
&= c^T H H^T c \\
&= c^T(\mu_1 P_1 + \mu_2 P_2 + \cdots + \mu_r P_r)c
\end{aligned}$$

$$
\begin{aligned}
&= c^T(\mu_1 P_1)c + c^T(\mu_2 P_2)c + \cdots + c^T(\mu_r P_r)c \\
&= \mu_1(c^T P_1 c) + \mu_2(c^T P_2 c) + \cdots + \mu_r(c^T P_r c) \\
&= \mu_1(c^T P_1 P_1 c) + \mu_2(c^T P_2 P_2 c) + \cdots + \mu_r(c^T P_r P_r c) \\
&= \mu_1(c^T P_1^T P_1 c) + \mu_2(c^T P_2^T P_2 c) + \cdots + \mu_r(c^T P_r^T P_r c) \\
&= \mu_1(P_1 c)^T(P_1 c) + \mu_2(P_2 c)^T(P_2 c) + \cdots + \mu_r(P_r c)^T(P_r c) \\
&= \mu_1 < P_1 c, P_1 c > + \mu_2 < P_2 c, P_2 c > + \cdots + \mu_r < P_r c, P_r c > \\
&= \mu_1 \|P_1 c\|^2 + \mu_2 \|P_2 c\|^2 + \cdots + \mu_r \|P_r c\|^2 \\
&= \mu_1 \|z_1\|^2 + \mu_2 \|z_2\|^2 + \cdots + \mu_r \|z_r\|^2 \\
&= \sum_{i=1}^{r} \mu_i \|z_i\|^2
\end{aligned}
$$

Entonces, $\|H^T c\|^2 = \sum_{i=1}^{r} \mu_i \|z_i\|^2$. Además,

$$
\begin{aligned}
\sum_{i=1}^{r} \mu_i \|z_i\|^2 &= \mu_1 \|z_1\|^2 + \sum_{i=2}^{r} \mu_i \|z_i\|^2 \\
&= \mu_1 \frac{w^2}{m} + \sum_{i=2}^{r} \mu_i \|z_i\|^2 \qquad \text{(por (4.22))}
\end{aligned} \tag{4.24}
$$

Observemos que se cumple lo siguiente usando las propiedades del Teorema espectral 3.3,

$$
\begin{aligned}
\sum_{i=1}^{r} \|z_i\|^2 &= \|z_1\|^2 + \|z_2\|^2 + \cdots + \|z_r\|^2 \\
&= \|P_1 c\|^2 + \|P_2 c\|^2 + \cdots + \|P_r c\|^2 \\
&= < P_1 c, P_1 c > + < P_2 c, P_2 c > + \cdots + < P_r c, P_r c > \\
&= (P_1 c)^T(P_1 c) + (P_2 c)^T(P_2 c) + \cdots + (P_r c)^T(P_r c) \\
&= c^T P_1^T P_1 c + c^T P_2^T P_2 c + \cdots + c^T P_r^T P_r c \\
&= c^T P_1 P_1 c + c^T P_2 P_2 c + \cdots + c^T P_r P_r c \\
&= c^T P_1^2 c + c^T P_2^2 c + \cdots + c^T P_r^2 c \\
&= c^T P_1 c + c^T P_2 c + \cdots + c^T P_r c \\
&= c^T (P_1 + P_2 + \cdots + P_r) c \\
&= c^T I c \\
&= c^T c \\
&= < c, c > \\
&= \|c\|^2
\end{aligned}
$$

Así, $\sum_{i=1}^{r} \|z_i\|^2 = \|c\|^2$, luego $\|z_1\|^2 + \sum_{i=2}^{r} \|z_i\|^2 = \|c\|^2$, entonces $\sum_{i=2}^{r} \|z_i\|^2 = \|c\|^2 - \|z_1\|^2$. Regresando a la ecuación (4.24),

$$
\begin{aligned}
\sum_{i=1}^{r} \mu_i \|z_i\|^2 &= \mu_1 \frac{w^2}{m} + \sum_{i=2}^{r} \mu_i \|z_i\|^2 \\
&\leqslant \mu_1 \frac{w^2}{m} + \mu_2 \sum_{i=2}^{r} \|z_i\|^2 \qquad \text{(pues } \mu_2 > \mu_3 > \cdots > \mu_r\text{)} \\
&= \mu_1 \frac{w^2}{m} + \mu_2(\|c\|^2 - \|z_1\|^2) \\
&= \mu_1 \frac{w^2}{m} + \mu_2(w - \frac{w^2}{m}) \qquad \text{(por (4.21) y (4.22))}
\end{aligned}
$$

Por lo tanto,

$$
\|H^T c\|^2 \leqslant \mu_1 \frac{w^2}{m} + \mu_2(w - \frac{w^2}{m}) \tag{4.25}
$$

Combinando las desigualdades (4.23) y (4.25) tenemos lo siguiente,

$$w(2d_x + d_c - 2) \leqslant \|H^T c\|^2 \leqslant \mu_1 \frac{w^2}{m} + \mu_2 (w - \frac{w^2}{m})$$

entonces por transitividad,

$$\begin{aligned} & w(2d_x + d_c - 2) \leqslant \mu_1 \frac{w^2}{m} + \mu_2 (w - \frac{w^2}{m}) \\ \Rightarrow \quad & 2d_x + d_c - 2 \leqslant \mu_1 \frac{w}{m} + \mu_2 (1 - \frac{w}{m}) \\ \Rightarrow \quad & (2d_x + d_c - 2)m \leqslant \mu_1 w + \mu_2 (m - w) \\ \Rightarrow \quad & (2d_x + d_c - 2)m - \mu_2 m \leqslant (\mu_1 - \mu_2) w \\ \Rightarrow \quad & \frac{(2d_x + d_c - 2 - \mu_2)m}{(\mu_1 - \mu_2)} \leqslant w \end{aligned}$$

Notemos que $dd_x \geqslant 2w$ ya que d es la distancia mínima o el peso mínimo de la palabra código x y w es el peso del vector c, entonces como

$$dd_x \geqslant 2w \qquad \text{y} \qquad w \geqslant \frac{m(2d_x + d_c - 2 - \mu_2)}{(\mu_1 - \mu_2)},$$

tenemos que,

$$d \geqslant \frac{2}{d_x} w \geqslant \left(\frac{2}{d_x}\right)\left(\frac{m(2d_x + d_c - 2 - \mu_2)}{(\mu_1 - \mu_2)}\right) \Rightarrow d \geqslant \frac{2m(2d_x + d_c - 2 - \mu_2)}{d_x(\mu_1 - \mu_2)}.$$

Ahora, notemos que dado que $\mathcal{B}$ es un grafo bipartito (d_x, d_c)-regular con n nodos variable y m nodos restricción, tenemos que el número de aristas incidentes a los nodos variable es nd_x y el número de aristas incidentes a los nodos restricción es md_c, entonces $nd_x = md_c$ eso implica que $d_x = \frac{md_c}{n}$. Así,

$$\begin{aligned} d \quad & \geqslant \quad \frac{2m(2d_x + d_c - 2 - \mu_2)}{d_x(\mu_1 - \mu_2)} \\ & = \quad \frac{2m(2d_x + d_c - 2 - \mu_2)}{\frac{md_c}{n}(\mu_1 - \mu_2)} \\ & = \quad \frac{2n(2d_x + d_c - 2 - \mu_2)}{d_c(\mu_1 - \mu_2)} \end{aligned}$$

Por lo tanto, la distancia mínima está acotada por $d \geqslant \frac{2n(2d_x + d_c - 2 - \mu_2)}{d_c(\mu_1 - \mu_2)}$. □

Veamos el siguiente ejemplo.

Ejemplo 4.5. *Sea $\mathcal{C}(\mathcal{B})$ el $[8,5]$-código LDPC $(2,4)$-regular con matriz de chequeo de paridad H (4.2) y grafo bipartito $\mathcal{B}$, Figura 46, con 8 nodos variable y 4 nodos restricción del Ejemplo 4.1. El objetivo es acotar la distancia mínima d del $[8,5]$-código LDPC $\mathcal{C}(\mathcal{B})$. De acuerdo al Teorema 4.2 para acotar la distancia mínima d del código LDPC $\mathcal{C}(\mathcal{B})$ necesitamos conocer el número de nodos variable n, el grado de los nodos variable d_x, el grado de los nodos restricción d_c, y el primer y segundo valor propio más grandes μ_1, μ_2 de la matriz HH^T que están asociados con los valores propios del grafo bipartito $\mathcal{B}$.*

Sabemos por el Ejemplo 4.1 que para dicho código LDPC $n = 8$, $d_x = 2$, $d_c = 4$, y por el Ejemplo 3.2 obtenemos en la salida (%o7) que los valores propios de la matriz HH^T son $\mu_1 = 8$ de multiplicidad $m_1 = 1$, $\mu_2 = 4$ de multiplicidad $m_2 = 1$, y $\mu_3 = 2$ con multiplicidad $m_3 = 2$, ordenados de tal forma que $\mu_1 > \mu_2 > \mu_3$. Así,

$$d \geqslant \frac{2n(2d_x + d_c - 2 - \mu_2)}{d_c(\mu_1 - \mu_2)} = \frac{2 \cdot 8(2 \cdot 2 + 4 - 2 - 4)}{4(8 - 4)} = \frac{16(2)}{16} = 2,$$

entonces la distancia mínima es mayor o igual a dos $d \geqslant 2$, esta cota nos proporciona mejor información que la anterior, y sabemos que en efecto esto se cumple, pues la distancia mínima es $d = 2$ para el $[8,5]$-código LDPC.

Comentario 4.2. *Estos dos teoremas probados por Tanner están generalizados para códigos LDPC irregulares, se pueden revisar tales resultados en [SKS05].*

4.5 EXPANSIÓN Y DISTANCIA MÍNIMA

En la sección 2.4 del capítulo 2 revisamos el concepto de expansión de un grafo bipartito. Y una buena expansión de un grafo implica buena propiedad de distancia mínima para el código detector-corrector de error asociado. El siguiente teorema muestra que dada una cota para uno de los parámetros de la expansión del grafo podemos dar una buena cota para la distancia mínima del código asociado.

Teorema 4.3. *Sea* $\mathcal{B} = (X \cup C, E)$ *una* (l, r, α, γ)*-expansión de un grafo bipartito con* $\gamma > \frac{1}{2}$*. Entonces el código* $\mathcal{C}(\mathcal{B})$ *asociado tiene distancia mínima* d *mayor que* αn*.*

Demostración. Procederemos por contradicción.

Sea $\mathcal{B} = (X \cup C, E)$ un grafo bipartito (l, r, α, γ)-expandido. Supongamos que $\gamma > \frac{1}{2}$ y que existe una palabra-código x diferente de cero de tamaño a lo más αn, así, su distancia mínima $d \leqslant \alpha n$. Sea $V \subseteq X$ que denota al conjunto de los nodos variable correspondientes a las posiciones no cero de la palabra-código x. Asumimos que cada nodo restricción vecino de V está conectado a V en un número par de veces. En particular, cada nodo restricción vecino debería estar conectado al menos dos nodos variable. Además, el número de aristas incidentes con V es $|V|l$ y como cada nodo restricción es incidente a al menos dos aristas, entonces el número de nodos restricción adyacentes a los nodos variable de la palabra-código x es a lo más $\frac{1}{2}|V|l$, esto es,

$$|\Gamma_V| \leqslant \frac{1}{2}|V|l \tag{4.26}$$

donde Γ_V denota al conjunto de nodos restricción adyacentes a los nodos de V.

Como el grafo bipartito es una (l, r, α, γ)-expansión se sigue que si $|V| \leqslant \alpha n$ entonces el número de nodos restricción vecinos es por lo menos $\gamma|V|l$, es decir,

$$|\Gamma_V| \geqslant \gamma|V|l. \tag{4.27}$$

Ahora dado que tenemos las siguientes desigualdades (4.26) y (4.27)

$$\frac{1}{2}|V|l \geqslant |\Gamma_V| \qquad \text{y} \qquad |\Gamma_V| \geqslant \gamma|V|l,$$

entonces,

$$\frac{1}{2}|V|l \geqslant \gamma|V|l \quad \Rightarrow \quad \frac{1}{2} \geqslant \gamma,$$

pero esto es una contradicción pues $\gamma > \frac{1}{2}$. Por lo tanto, el código $\mathcal{C}(\mathcal{B})$ asociado al grafo bipartito $\mathcal{B}$ tiene distancia mínima $d > \alpha n$. □

Comentario 4.3. *Un concepto especial a partir de conocer los grafos expandidos son los códigos expandidos, los cuales pertenecen a la clase de los códigos LDPC, de los cuales podemos encontrar varios trabajos relacionados con estos conceptos y las cotas para la distancia mínima, algunas referencias de consulta son [Kelo9] y [SS96].*

Lo importante de cada uno de estos teoremas que hemos revisado es que al dar una cota sobre la distancia mínima es posible aproximar cuantos errores es capaz de corregir un código, pues sabemos que la distancia mínima ayuda a determinar el número de errores que puede corregir un código, ver el Teorema 2.6, en particular un código LDPC.

BIBLIOGRAFÍA

[ADH98] A. S. Asratian, T. M. J. Denley, and R. Häggkvist, *Bipartite graphs and their applicatons*, first ed., Cambridge University Press, 1998.

[AGM87] N. Alon, Z. Galil, and V. D. Milman, *Better expanders and superconcentrators*, Journal of Algorithms (1987), no. 8, 337–347.

[Alo88] N. Alon, *Eigenvalues and expanders*, Combinatorica **6** (1988), no. 2, 83–96.

[BC09] R. A. Brualdi and D. Cvetković, *A combinatorial approach to matrix theory and its applications*, Taylor & Francis Group, 2009.

[BH12] A. E. Brouwer and W. H. Haemers, *Spectra of graphs*, Springer, 2012.

[Big74] N. Biggs, *Algebraic graph theory*, Cambridge University Press, 1974.

[BM08] J. A. Bondy and U. S. R. Murty, *Graphs theory*, Springer, 2008.

[CBFS08] J. Cannon, W. Bosma, C. Fieker, and A. Steel, *Magma Computer Algebra System*, (Versión 2.14-15) [Software] (2008), Disponible en http://magma.maths.usyd.edu.au/magma/.

[CDS80] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of graphs, theory and applications*, Academic Press, 1980.

[CRS97] D. Cvetković, P. Rowlinson, and S. Simić, *Eigenspaces of graphs*, first ed., Cambridge University Press, 1997.

[CRS10] ———, *An introduction to the theory of graph spectra*, first ed., Cambridge University Press, 2010.

[Die05] R. Diestel, *Graph theory*, third ed., Springer-Verlag Heidelberg, 2005.

[FR01] B. Fry and C. Reas, *Processing 2*, (Versión 2.2.1) [Software] (2001), Disponible en https://processing.org/.

[FZ11] A. Frolov and V. Zyablov, *Upper and lower bounds on the minimum distance of expander codes*, IEEE International Symposium on Information Theory Proceedings (2011), 1397–1401.

[Gal63] R. G. Gallager, *Low-Density Parity-Check Codes*, Ph.D. thesis, Cambridge, MA: MIT Press, 1963.

[GG81] O. Gabber and Z. Galil, *Explicit constructions of linear-sized superconcentrators*, Computer and System Sciences (1981), no. 22, 407–420.

[Gri04] R. P. Grimaldi, *Discrete and combinatorial mathematics an applied introduction*, fifth ed., Pearson Addison Wesley, 2004.

[HK71] K. Hoffman and R. Kunze, *Linear algebra*, second ed., Prentice-Hall, Inc., 1971.

[HLW06] S. Hoory, N. Linial, and A. Wigderson, *Expander graphs and their applications*, Bulletin of the American Mathematical Society **43** (2006), no. 4, 439–561.

[Kel09] C. A. Kelley, *Minimum distance and pseudodistance lower bounds for generalized LDPC codes*, Int. J. Inform. and Coding Theory (2009), 1–23.

[LA14] C. A. López-Andrade, *Matemáticas y sus aplicaciones 4*, primera ed., ch. 1, pp. 5–33, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, México, 2014.

[Lhé04] A. Lhéritier, *Códigos de Paridad de Baja Densidad (Low-Density Parity-Check, LDPC)*, Tech. report, Universidad de la República, 2004.

[Lub10] A. Lubotzky, *Discrete gropus, expanding graphs and invariant measures*, Birkhäuser Verlag, 2010.

[Mac03] D. J. C. Mackay, *Information theory, inference, and learning algorithms*, 2003.

[McE04] R. J. McEliece, *The theory of information and coding*, second ed., Cambridge University Press, 2004.

[MM64] M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Inc., 1964.

[Poo11] D. Poole, *Álgebra lineal, una introducción moderna*, tercera ed., Cengace Learning, 2011.

[Rom92] S. Roman, *Coding and information theory*, Springer-Verlag, 1992.

[RU01] T. J. Richardson and R. L. Urbanke, *The capacity of Low-Density Parity-Check codes under message-passing decoding*, IEEE Trans. Inform. Theory **47** (2001), no. 2, 599–618.

[RU08] T. Richardson and R. Urbanke, *Modern coding theory*, first ed., Cambridge University Press, 2008.

[Sha48] C. E. Shannon, *A mathematical theory of communication*, The Bell System Technical Journal **27** (1948), 379–423,623–656.

[SKS05] M-H. Shin, J-S. Kim, and H-Y. Song, *Generalization of Tanner's minimum distance bounds for LDPC codes*, IEEE Communications Letters **9** (2005), no. 3, 240–242.

[SS96] M. Sipser and D. A. Spielman, *Expander codes*, IEEE Trans. Inform. Theory **42** (1996), no. 6, 1710–1711.

[Tan01] R. M. Tanner, *Minimum-distance bounds by graph analysis*, IEEE Trans. Inform. Theory **47** (2001), no. 2, 888–889.

[VLIK15] A. Vodopivec, Z. Lenarcic, D. Ilijev, and G. Königsmann, *wxMaxima 16.04.2*, (Versión 5.38.1) [Software] (2015), Disponible en http://maxima.sourceforge.net/es/.

[WK03] S. B. Wicker and S. Kim, *Fundamentals of codes, graphs, and iterative decoding*, Kluwer Academic Publishers, 2003.