Universidad Nacional Autónoma
de México

Posgrado en Ciencia e Ingeniería
de la Computación

COTAS PARA CÓDIGOS SOBRE GRUPOS CUÁNTICOS FINITOS

T E S I S

QUE PARA OBTENER EL GRADO DE:
MAESTRA EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA:
MAYRA LORENA DÍAZ SOSA

DIRECTOR DE TESIS:
DR. VLADISLAV KHARTCHENKO

DICIEMBRE DE 2008.

RESUMEN

El propósito de este trabajo es encontrar y comparar algunas cotas para un código cuyo alfabeto sea construido sobre grupos cuánticos finitos. Concretamente, se analizan las cotas asintóticas de Hamming, Elias y Gilbert-Varshamov. Dentro de la teoría de grupos cuánticos finitos la tesis se enfoca al caso de los anillos de Frobenius finitos que cuentan con pesos homogéneos.

Se presentan los fundamentos de la teoría de códigos detectores de errores y se explica el problema principal en la teoría de códigos. Se exponen las cotas superiores e inferiores más destacadas de la teoría de códigos clásica y se detalla su cálculo. Asimismo se explica la relación de los anillos de Frobenius finitos con el grupo cuántico pequeño de Luzstig. Se hace una revisión detallada del artículo de Greferath y O'Sullivan (2004) y se exponen las cotas de Plotkin y de Elias para códigos sobre anillos de Frobenius finitos con pesos homogéneos. Finalmente, se explican las versiones asintóticas de las cotas de Hamming, Elias y Gilbert-Varshamov sobre los citados anillos.

Palabras clave: códigos correctores de errores, códigos sobre anillos finitos, pesos homogéneos, cotas superiores, cotas asintóticas, anillos de Frobenius finitos.

Tesis de Maestría apoyada por el Consejo Nacional de Ciencia y Tecnología (CONACyT) de México, CVU 205912.

PREFACIO

Durante mis estudios de licenciatura en Actuaría en la Facultad de Estudios Superiores Acatlán tuve oportunidad de cursar algunas materias sobre matemáticas abstractas como son álgebra lineal, álgebra superior y geometría analítica, entre otras. En ocasiones me preguntaba si estas ramas de las matemáticas encontraban aplicaciones prácticas en la vida cotidiana. Esta inquietud, aunada a mi inclinación por la programación, me llevó a realizar mis estudios de maestría en el Posgrado en Ciencia e Ingeniería de la Computación en la máxima casa de estudios.

La teoría de códigos es el resultado de la maravillosa combinación de la teoría de códigos correctores de errores y matemáticas para la modelación de la comunicación confiable en la presencia de ruido, involucrando matemáticas discretas, cálculo combinatorio, álgebra moderna, álgebra lineal, teoría de probabilidad y estadística. La teoría de códigos ha sido investigada y desarrollada durante más de cinco décadas y ha visto gran aplicación en diversos ámbitos que involucran la transmisión de información codificada, de ahí mi inclinación para elegir este tema de tesis. Mientras que originalmente la teoría algebraica de códigos correctores de errores tuvo lugar en el escenario de los espacios vectoriales sobre campos finitos, el estudio de los códigos lineales sobre anillos finitos ha cobrado fuerza e importancia a partir de que, años atrás, especialistas en la materia descubrieron que códigos aparentemente no lineales en realidad están relacionados con códigos lineales sobre el anillo de los enteros módulo cuatro (véanse los artículos de Calderbank y otros (1993), Conway y Sloane (1993), Pless y otros (1997) y Pless y Qian (1996)).

Motivados por esos estudios, los matemáticos se han propuesto encontrar cotas para el estudio de códigos sobre anillos finitos. Una pregunta natural en esta búsqueda es si los componentes esenciales de la teoría clásica (resultados como los teoremas de equivalencia de MacWilliams o las identidades de MacWilliams) se cumplen para un código sobre un anillo en particular. Por supuesto, la respuesta a esta pregunta consta de una belleza intrínseca propia de las matemáticas abstractas, pero la importancia de sus aplicaciones potenciales no debe ser subestimada. De acuerdo con Honold (2001), comprender cuándo un tipo de anillo satisface alguna versión del teorema de equivalencia de MacWilliams, por ejemplo, asegura a los desarrolladores de códigos que, tan pronto como se adopten dichos anillos, la singularidad de sus resultados se convertirá en algo sencillo de verificar puesto que la existencia de códigos equivalentes los reducirá a

casos más tratables. Ello resulta una tarea de vital importancia, por ejemplo, cuando se trata de patentar un sistema de codificación. Dos artículos de Wood [(1999), (2006)] muestran que si la extensión del teorema de MacWilliams se cumple para un anillo finito, entonces el anillo debe ser de Frobenius, resultado que se cumple también a la inversa. Esto es una fuerte evidencia de que los anillos de Frobenius son los anillos más apropiados para la teoría de códigos. Por esta razón el problema de la construcción de anillos de Frobenius finitos ha cobrado vital importancia. Afortunadamente, la teoría de los grupos cuánticos, una rama del álgebra moderna, proporciona ejemplos de álgebras de Frobenius.

Para la realización de la tesis se investigó y recabó información sobre la teoría de campos finitos y sus aplicaciones en la teoría de códigos, estructuras algebraicas finitas no estándares en la teoría de códigos, y cotas calculadas en años recientes para códigos sobre anillos de Frobenius finitos. Por otra parte, en la sección de anexos se incluyen, como parte de mis actividades de posgrado para realizar esta investigación, las constancias de participación en el *II Taller de álgebra y topología* realizado en la Facultad de Ciencias de la Universidad Autónoma del Estado de Morelos, y en la *Escuela de Modelación y Métodos Numéricos*, misma que versó sobre supercómputo y sus aplicaciones y que tuvo lugar en el Centro de Investigación en Matemáticas, en Guanajuato.

Existen diversos lineamientos y herramientas potencialmente útiles para la realización de trabajos de investigación, la publicación de artículos y la elaboración de tesis que debiesen resultar conocidos para los alumnos de posgrado, e incluso de licenciatura. Sin embargo, deseo hacer una breve mención de algunos de los lineamientos y herramientas empleados en esta tesis con la intención de que esta información le resulte útil a quien consulte este trabajo:

(i) Este trabajo de tesis fue estructurado y elaborado conforme a los estándares internacionales ISO 7144 para la documentación y presentación de tesis y documentos similares, mismos que son vigentes a la fecha.

(ii) En general, los trabajos de investigación requieren de consultar literatura actual, pertinente, relevante y vigente que suele tener un costo significativo. El acceso gratuito a los artículos citados en la tesis fue posible a través de la Biblioteca Digital de la Universidad Nacional Autónoma de México.

(iii) Además, identificar información relevante y vigente no siempre es tarea fácil. ISI Web of Knowledge es una poderosa plataforma de investigación en constante actualización que ayuda a analizar, encontrar de forma rápida y compartir información sobre diversos ámbitos del conocimiento. Cuenta con una base de datos sumamente amplia en diversas categorías (artículos, sitios web, memorias, patentes, libros, etc.) y con una herramienta de análisis que determina los autores más

prolíficos y citados por tema, proporciona un panorama general sobre las tendencias en investigación en diversos ámbitos e identifica proyectos de investigación prominentes por su vigencia y relevancia.

(iv) La escritura de textos científicos frecuentemente involucra el uso de símbolos y fórmulas matemáticas que implican una ardua labor en procesadores de texto convencionales. La tesis fue escrita en LaTeX 2_ε, el cual ofrece una gran calidad tipográfica para textos científicos razón por la que las más importantes casas editoriales, como Springer-Verlag, lo emplean y los congresos internacionales y revistas científicas arbitradas lo exigen en la presentación de proyectos.

Por último, no quisiera concluir sin mencionar que a lo largo de estos dos años he contado con diversos apoyos sin los cuales no habría sido posible la culminación de este trabajo. Por esta razón deseo manifestar mi profundo agradecimiento:

A mis padres, María Concepción Sosa Tuñón y Raúl Díaz Nájera, por su cariño y apoyo.

A Víctor M. Rangel Cortés, por siempre creer en mí, por apoyarme e impulsarme.

A la Universidad Nacional Autónoma de México, mi casa de estudios a la que tanto debo; a la Facultad de Estudios Superiores Cuautitlán, al Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas, y a todos mis profesores del Posgrado. En especial deseo agradecerle a mi tutor y director de tesis, el Dr. Vladislav Khartchenko, su dedicación, todo su apoyo, su paciencia y por su gentileza al haberme nombrado su ayudante; por convertirse en mi diablo de los números y convertirme en su iniciada... Balchói spasiba na vsigdá!

A la Dra. MariCarmen González Videgaray y al M. en C. Rubén Romero Ruiz por compartirme sus conocimientos, guiar mis pasos y enseñarme siempre con tanto cariño y paciencia.

Al Consejo Nacional de Ciencia y Tecnología, por el valioso apoyo que me proporcionó durante mis estudios de maestría.

A todos mis amigos del posgrado, por su cariño y por haber compartido conmigo estos dos años.

A mis alumnos, quienes llenan mi vida de una manera que ni siquiera sospechan; por motivarme y hacerme comprender que la docencia le da sentido a mi existencia.

Mayra Lorena Díaz Sosa.

Diciembre de 2008.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

INTRODUCCIÓN

En 1948, Claude Shannon en su artículo *A Mathematical Theory of Communication* mostró que dado un canal de comunicación con ruido existe un número, llamado la capacidad del canal, tal que es posible lograr una comunicación confiable a cualquier tasa por debajo de la capacidad del canal si se emplean técnicas de codificación y decodificación adecuados. La publicación de dicho artículo marca el nacimiento de la teoría de códigos, un área de las matemáticas aplicadas dedicada al estudio de la transmisión de datos por medio de un canal con ruido y la recuperación de mensajes corrompidos. Desde entonces la teoría de códigos ha experimentado un crecimiento sorprendente y ha encontrado aplicaciones diversas en los sistemas de comunicación, en la industria de los discos compactos y en dispositivos de almacenamiento, por citar sólo algunos ejemplos.

La siguiente es la definición del problema principal en la teoría de códigos clásica. $C_q(n, M, d)$ denota un código definido sobre F_q, de longitud n, tamaño M y distancia d. Considerando un valor de n fijo el tamaño M del código es un parámetro que permite medir la eficiencia del mismo (en términos de la tasa de transmisión) y la distancia d indica la capacidad para corregir errores del código en cuestión. Por supuesto, sería deseable que un código fuese eficiente y que a la vez permitiera corregir el mayor número de errores posible, esto es, que tanto M como d fueran tan grandes como fuese posible. Sin embargo, esto no ocurre, pues la eficiencia y la capacidad de corregir errores de un código están estrechamente ligadas. El problema de determinar el mayor valor posible de M para el que existe un código (n, M, d) sobre un alfabeto A de tamaño $q > 1$ es comúnmente conocido como el problema principal en teoría de códigos. Se dice que el código con parámetros (n, M, d) es óptimo cuando su tamaño M es el máximo posible.

Mientras que es difícil determinar el valor exacto de M para el cual un código es óptimo dados los parámetros n y d, existen diferentes cotas superiores que permiten ubicar en qué rango de valores enteros se encuentra M. Entre las cotas superiores más famosas se encuentran las de Hamming, de Singleton y de Plotkin. Ésta última destaca entre las anteriores por ofrecer un margen más pequeño de valores para M (en algunos casos supera por mucho a las otras citadas). Este problema también se presenta bajo un enfoque cuántico al emplearse grupos cuánticos finitos para construir un alfabeto A en lugar de construirlo sobre campos finitos, y éste es el objeto de estudio de la presente investigación.

El desarrollo moderno de la teoría de códigos correctores de errores incluye la consideración de los códigos sobre objetos más complejos que los campos finitos. De acuerdo con Honold (2001), desde que Nakayama introdujo la noción de anillo de Frobenius se ha venido desarrollando una rica teoría sobre anillos de Frobenius. Hasta hace poco se descubrió que los anillos de Frobenius finitos encuentran aplicación práctica en la teoría de códigos. Klemm (1989) demostró las identidades de MacWilliams sobre anillos de Frobenius finitos y conmutativos, mientras que Nechaev y otros (1997) desarrollaron una teoría general de códigos lineales sobre R-módulos de anillos finitos conmutativos, incluyendo su dualidad y las identidades de MacWilliams. Wood (1999) probó también las identidades de MacWilliams y la extensión del teorema de MacWilliams para anillos de Frobenius finitos arbitrarios. Concretamente, empleando argumentos teóricos sobre caracteres, Wood (1999) probó que toda equivalencia entre códigos sobre anillos de Frobenius finitos podía establecerse por medio de una transformación monomial. Después, Greferath y Schmidt (2000) realizaron otra prueba de este mismo resultado empleando métodos combinatorios. La importancia de este resultado radica en el hecho de que, hasta la fecha, la mayoría de los anillos finitos que han sido considerados como alfabetos potenciales para los códigos correctores de errores han resultado ser anillos de Frobenius. Por esta razón el problema de la construcción de anillos de Frobenius finitos ha cobrado cada vez mayor importancia.

Los grupos cuánticos son ejemplos particulares de álgebras de Hopf. El teorema de Larson y Sweedler (1969) prueba que toda álgebra de Hopf de dimensión finita en realidad es de Frobenius, mientras que un estudio reciente de Skryabin (2007) señala que cada subálgebra de coidealo de un sólo lado también es un álgebra de Frobenius. Estos teoremas proporcionan ejemplos importantes de anillos de Frobenius finitos si consideramos álgebras de Hopf y sus subálgebras con coidealo derecho sobre un campo finito F_q con $q = p^n$ elementos (donde p es un número primo).

El propósito de este trabajo es encontrar y comparar algunas cotas para un código cuyo alfabeto sea construido sobre grupos cuánticos finitos. Concretamente, se analizan las cotas asintóticas de Hamming, Elias y Gilbert-Varshamov. Dentro de la teoría de grupos cuánticos finitos la tesis se enfoca al caso de los anillos de Frobenius finitos que cuentan con pesos homogéneos, concepto que se debe a Constantinescu y Heise (1997).

En el Capítulo 1, *Nociones fundamentales de la Teoría de Códigos*, se hace una reseña de los orígenes de la teoría de códigos. Se explican los parámetros de un código así como conceptos fundamentales de la misma y cómo aplicar los conceptos de distancia mínima de un código y de la distancia de Hamming en la detección y corrección de un código. Se aborda el caso de los códigos lineales y su utilidad para la codificación y decodificación de información por distintos métodos. A lo largo de este capítulo se incluyen múltiples ejemplos que ilustran los conceptos abordados.

El Capítulo 2, *Cotas en Teoría de Códigos*, explica de manera formal en qué consiste el problema principal en la Teoría de Códigos y describe dos cotas inferiores: las cotas de la esfera y de Gilbert-Varshamov; y cuatro cotas superiores, las cotas de Hamming, de Singleton, de Plotkin y la de Griesmer. Asimismo, se detallan los pasos en la deducción de dichas cotas.

En el Capítulo 3, *Anillos de Frobenius finitos relacionados al grupo cuántico de Lusztig*, se hace un breve repaso de la forma polar de los números complejos con la finalidad de ilustrar claramente el concepto de caracter. Se explican la construcción del carácter principal de un álgebra de Frobenius finita y los anillos de Frobenius finitos relacionados al grupo cuántico de Lusztig. Además, se describe un algoritmo que permite encontrar una representación en la forma de Lusztig para cualquier producto de elementos base, lo que a su vez permite calcular la forma de los elementos del anillo y encontrar su caracterización.

En el Capítulo 4, *Cotas para códigos sobre anillos de Frobenius y de pesos homogéneos*, se define el concepto de peso homogéneo y se desarrollan dos cotas superiores para códigos de peso homogéneo y construidos sobre anillos de Frobenius: la cota de Plotkin y la Cota de Elias. Se desarrolla también la versión asintótica de la cota de Elias y se hace un comparativo entre las cotas de Elias, de Gilbert-Varshamov y de Hamming, todas en sus versiones asintóticas.

Finalmente, se presentan las conclusiones.

1. NOCIONES FUNDAMENTALES DE LA TEORÍA DE CÓDIGOS

-"Minino de Cheshire", comenzó ella algo tímidamente,
"me dirás por favor, ¿qué camino debo tomar para irme de aquí?"
-"Eso depende mucho de dónde quieras ir", dijo el gato.
-"Poco me preocupa dónde ir", contestó Alicia.
-"Entonces nada importa qué camino tomes", replicó el gato.
Lewis Carroll, Alice in wonderland.

Supongamos que un día llegamos a casa y encontramos un mensaje escrito que dice: "¡Perdí pXso!", en donde X significa que no podemos identificar la letra que fue escrita entre la p y la s. Podrían pasar por nuestra mente dos posibilidades para el mensaje: "¡Perdí peso!" o "¡Perdí piso!". ¡Una sóla letra cambiaría por completo el sentido de la oración! ¿Cuál es el mensaje que deseaban comunicarnos? Esta es la esencia de la teoría de códigos.

1.1. Antecedentes

El término *teoría de códigos* se utiliza con frecuencia para referirse en realidad a dos áreas cuya relación es estrecha: la teoría de los códigos correctores de errores y las matemáticas empleadas para modelar la comunicación fiable en la presencia de ruido. Los códigos correctores de errores son colecciones de secuencias de elementos de un conjunto finito, normalmente palabras sobre un alfabeto finito, con algunas propiedades. La teoría de los códigos correctores de errores pertenece al área de las matemáticas combinatorias. Por otra parte, la teoría de la comunicación en presencia de ruido pertenece a las áreas de teoría de la información y/o estadística. Estos dos aspectos de la teoría de códigos gozan de una relación muy cercana desde sus orígenes.

La teoría de códigos vio sus inicios con los trabajos de R. Hamming y C. Shannon a finales de la década de 1940. Por una parte, Hamming estudiaba dispositivos para el almacenamiento de información con el propósito de encontrar algún método que sirviera para evitar la corrupción de la misma en forma de bits. Se dio cuenta de la necesidad

inminente de considerar conjuntos de bits, palabras o mejor dicho palabras código, en los que hubiera parejas que difirieran entre sí en un número determinado de coordenadas. Hamming definió a partir de esta observación una interesante métrica, a nuestros días mejor conocida como distancia de Hamming, para cuantificar la distancia entre dos palabras código, métrica que cuenta con una serie de propiedades muy interesantes; también aportó una familia de códigos cuya distancia es no trivial. La distancia y los códigos de Hamming serán estudiados a mayor detalle a lo largo del presente trabajo. Cabe destacar que las publicaciones de Hamming fueron en 1950.

Poco antes de que se publicara el trabajo de Hamming, Shannon (1948) escribió un tratado en el que formalizaba, matemáticamente hablando, la teoría de la comunicación. En dicho trabajo, Shannon, estudió cómo un emisor podría enviar de forma eficiente información a un receptor por medio de un canal de comunicación. Planteó dos vertientes del canal de comunicación: sin ruido y con ruido. En el caso del canal sin ruido, la idea era comprimir la información para que el emisor la enviara con el menor número posible de símbolos (en este caso bits) a la vez que el receptor fuera capaz de recuperar dicha información correctamente. El canal con ruido, el caso de mayor interés en la teoría de códigos, es aquel en el cual el canal de comunicación altera la señal enviada añadiéndole "ruido", es decir, distorsiona la información. En este caso la idea es añadir redundancia a la información, antes de ser enviada, que permita al receptor ubicar algunos pocos símbolos o bits incorrectos para que pueda recuperar la información enviada originalmente por el emisor. A lo largo de estas investigaciones, Shannon descubrió que su trabajo servía también para establecer la tasa de transmisión de la información en diferentes tipos de canales, y también descubrió que cuando se transmite información a una tasa fija la recuperación de la información de "mensajes" largos es mucho más sencilla y factible que la de los cortos.[1]

Prácticamente, los trabajos de Hamming y de Shannon coinciden cronológicamente, pero además se complementan el uno al otro de una manera sorprendente. Hamming se concentró en aspectos combinatorios; Shannon, en modelos probabilísticos y en un enfoque estadístico.

De acuerdo con Sudan (2001), aunque el artículo de Shannon apareció primero, éste tomó un código propio de Hamming citando a Hamming, por lo que se sabe que estaba consciente del trabajo realizado hasta ese momento por Hamming. Aunque Hamming no citó a Shannon de forma explícita en su trabajo, citó a Golay, quien por su parte citó a Shannon, por lo que es posible que, al momento de publicar su artículo, Hamming también estuviera al tanto del trabajo desarrollado por Shannon.

Como sea, los trabajos de Shannon y Hamming tuvieron gran impacto y han dado pie

[1] Entiéndase la palabra mensaje en el sentido más amplio posible, ya que puede tratarse de una imagen, una canción, una frase, etc.

a múltiples investigaciones en materia de la teoría de la información y de las ciencias de la computación.

Antes de comenzar a revisar algunos de los aspectos más sobresalientes en materia de teoría de códigos, es importante contar con conocimientos previos en materia de matemáticas discretas, combinatoria, álgebra lineal y álgebra moderna, por lo que se recomiendan los libros de Grimaldi (1994), Birkoff y Mac˜Lane (1965), Herstein (1975), y Lay (1997) para hacer un repaso general en estas ramas de las matemáticas.

1.2. Códigos, distancia y peso

Como se ha mencionado existen tres entidades involucradas en el proceso de transmitir información: el emisor, el receptor y un canal con ruido. El objetivo del emisor es enviar un mensaje m de un conjunto finito $\mathcal{M}$, el espacio de mensajes, al receptor. Tanto el emisor como el receptor conocen el espacio de mensajes posibles, mismo que usualmente tiene un tamaño grande.[2]

El canal con ruido puede comunicar arbitrariamente secuencias largas de símbolos de un alfabeto determinado, digamos A. Para enviar mensajes por este canal, el emisor y el receptor deben estar de acuerdo con la longitud de las cadenas de símbolos que se van a enviar, longitud frecuentemente denotada por n. Por esta razón, el espacio de palabras que se transmiten por el canal es de A^n.

Puesto que la idea es la de compartir información, tanto el emisor como el receptor deben conocer la codificación E, el mapeo inyectivo (uno a uno)

$$E : \mathcal{M} \to A^n$$

que se utilizará para codificar la información antes de ser transmitida, de manera que el emisor enviará por el canal la información como $\mathbf{c} = E(m)$ en lugar de m. La imagen de la función $\{\mathbf{c}|m \in \mathcal{M}\}$ es lo que llamamos un *código corrector de errores* o, simplemente, una *código*.

Cuando el mensaje "pasa" por el canal éste último puede añadirle ruido o errores, es decir, puede modificar algunos de los símbolos de las palabras código transmitidas en cuyo caso se tiene un error $\mathbf{e} \in A^n$ y entonces el mensaje que le llega al receptor es $\mathbf{x} = E(m) + \mathbf{e}$. A continuación, el receptor emplea una función decodificadora tal que

$$D : A^n \to \mathcal{M}.$$

[2] Note que en este momento no aparece entre el emisor y el receptor el concepto de canal de comunicación todavía.

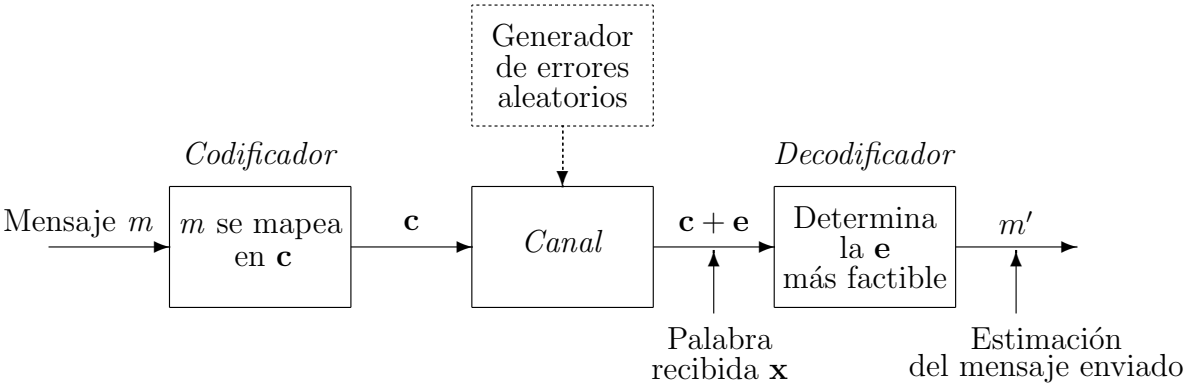


Fig. 1.1: Canal con ruido

Por supuesto, lo ideal sería que cuando el receptor decodificara la información obtuviera $D(\mathbf{x}) = m$, esto es, el mensaje que el emisor deseaba enviarle originalmente.

Es importante destacar algunas de las características deseables de los códigos:

(i) rápida codificación de mensajes,

(ii) fácil transmisión de los mensajes codificados,

(iii) fácil decodificación de los mensajes recibidos,

(iv) máxima transferencia de información por unidad de tiempo, y

(v) máxima capacidad para detectar y/o corregir errores posible.

Como veremos de aquí en adelante, el álgebra moderna y el álgebra lineal, son las ramas de las matemáticas sobre las que se basa la teoría de códigos, principalmente.

1.1 Definición. Sea $A = \{a_1, a_2, \ldots, a_q\}$ un conjunto de tamaño q al que llamaremos *alfabeto* del código y cuyos elementos a_i para $1 \leq i \leq n$ son llamados *símbolos* del código.

(i) La *palabra q-aria* de longitud n sobre A es una secuencia $\mathbf{w} = w_1 w_2 \ldots w_n$, donde cada $w_i \in A$ para toda i. [3]

(ii) Un *código de bloque q-ario* de longitud n sobre A es un conjunto no vacío C de palabras q-arias con la misma longitud n.

[3] $\mathbf{w}$ puede considerarse como un vector $(w_1 \ldots w_n)$.

(iii) Cada elemento de C es llamado *palabra código* en C.

(iv) Al número de palabras código en C, denotado por $|C|$, se le llama *tamaño* de C.

(v) La *tasa de información* del código C de longitud n se define como $(\log_q |C|)/n$.

(vi) Un código de longitud n y tamaño M se llama *código*-(n, M).

Normalmente, el alfabeto del código se toma como un campo finito F_q de orden q.

Ahora, definiremos formalmente el concepto de canal.

1.2 Definición. Un *canal de comunicación* consiste en un alfabeto $A = \{a_1, a_2, \ldots, a_q\}$ y un conjunto de probabilidades $\mathcal{P}(a_j \text{ recibida}|a_i \text{ enviada})$, tales que

$$\sum_{j=1}^{q} \mathcal{P}(a_j \text{ recibida}|a_i \text{ enviada}) = 1,$$

para toda i, donde $\mathcal{P}(a_j \text{ recibida}|a_i \text{ enviada})$ es la probabilidad condicional de que se reciba el símbolo a_j siendo que el símbolo a_i fue enviado.

1.3 Definición. Se dice que un canal de comunicación no tiene *memoria* si el resultado de cualquier transmisión es independiente del resultado en transmisiones previas, i.e., si $\mathbf{c} = c_1c_2 \ldots c_n$ y $\mathbf{x} = x_1x_2 \ldots x_n$ son palabras de longitud n, entonces

$$\mathcal{P}(\mathbf{x} \text{ recibida}|\mathbf{c} \text{ enviada}) = \prod_{i=1}^{n} \mathcal{P}(x_i \text{ recibida}|c_i \text{ enviada}).$$

1.4 Definición. Un canal q-ario *simétrico* es un canal sin memoria con alfabeto de tamaño q tal que:

(i) Cada símbolo transmitido tiene la misma probabilidad $p < 0.5$ de ser recibido por error.

(ii) Si un símbolo se recibe por error, entonces cada uno de los posibles errores en la transmisión de los $q-1$ símbolos restantes tiene la misma probabilidad de ocurrir.

Cabe destacar que, en la práctica, es deseable que la probabilidad de error en la transmisión p sea mucho menor a 0.5.

Ahora que hemos hablado del error en la transmisión de información estamos listos para introducir dos nuevos conceptos particularmente útiles para el estudio del mismo: la distancia y el peso de un código.

1.5 Definición. Sean $\mathbf{x}$ e $\mathbf{y}$ palabras de longitud n sobre el alfabeto A. La *distancia* (*de Hamming*) de $\mathbf{x}$ a $\mathbf{y}$, denotada por $d(\mathbf{x}, \mathbf{y})$, se define como el número de *coordenadas* en las que $\mathbf{x}$ e $\mathbf{y}$ difieren. Si $\mathbf{x} = x_1 \dots x_n$ e $\mathbf{y} = y_1 \dots y_n$, entonces

$$d(\mathbf{x}, \mathbf{y}) = d(x_1, y_1) + \dots + d(x_n, y_n), \tag{1.1}$$

donde x_i e y_i son las i-ésimas coordenadas (palabras de longitud 1), y

$$d(x_i, y_i) = \begin{cases} 1 & \text{si } x_i \neq y_i; \\ 0 & \text{si } x_i = y_i. \end{cases}$$

1.6 Proposición. *Sean* $\mathbf{x}$, $\mathbf{y}$ *y* $\mathbf{z}$ *palabras de longitud* n *sobre* A. *La distancia de Hamming satisface las siguientes propiedades:*

(i) $0 \leq d(\mathbf{x}, \mathbf{y}) \leq n$,

(ii) $d(\mathbf{x}, \mathbf{y}) = 0$, *si y sólo si* $\mathbf{x} = \mathbf{y}$,

(iii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$,

(iv) $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ *(desigualdad triangular).*

Demostración. (i), (ii) y (iii) se siguen directamente de la definición de distancia de Hamming.

Por (1.5), basta con probar (iv) para $n = 1$, en cuyo caso existen dos posibilidades:

Si $\mathbf{x} = \mathbf{z}$, entonces (iv) se cumple puesto que $d(\mathbf{x}, \mathbf{z}) = 0$.
Si $\mathbf{x} \neq \mathbf{z}$, entonces o $\mathbf{y} \neq \mathbf{x}$ o $\mathbf{y} \neq \mathbf{z}$, por lo que se cumple nuevamente (iv).

□

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

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

1.8 Definición. Sea $\mathbf{x}$ una palabra en F_q^n. El *peso* (de Hamming) de $\mathbf{x}$, denotado por $wt(\mathbf{x})$, es el número de coordenadas o símbolos distintos de cero en $\mathbf{x}$; es decir,

$$wt(\mathbf{x}) = d(\mathbf{x}, \mathbf{0}),$$

donde $\mathbf{0}$ es la palabra cero.

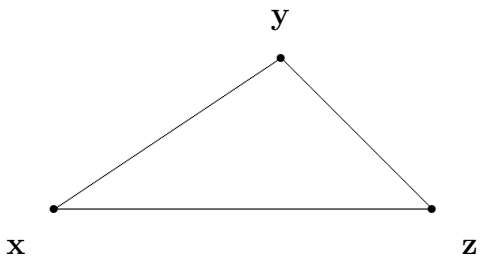


Fig. 1.2: Desigualdad triangular en la distancia de Hamming: $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

Para cada elemento x de F_q, el peso de Hamming es simplemente:

$$wt(x) = d(x, 0) = \begin{cases} 1 & \text{si } x \neq 0; \\ 0 & \text{si } x = 0. \end{cases}$$

Al igual que en el caso de la distancia de Hamming, el peso de Hamming puede definirse de forma equivalente como

$$wt(\mathbf{x}) = wt(x_1) + \ldots + wt(x_n) \tag{1.2}$$

donde $\mathbf{x} \in F_q^n$.

A continuación revisaremos algunas de las propiedades más importantes sobre el peso de Hamming.

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

Demostración. Para $x, y \in F_q$, $d(x, y) = 0$ si y sólo si $x = y$, lo cual es cierto si y sólo si $x - y = 0$ o, equivalentemente, $wt(x - y) = 0$. El lema se sigue de (1.1) y (1.2). □

Una consecuencia inmediata del Lema 1.9 es el siguiente corolario, en virtud de que siempre que q sea par $a = -a$ para toda $a \in F_q$.

1.10 Corolario. *Sea* q *par. Si* $\mathbf{x}, \mathbf{y} \in F_q^n$*, entonces* $d(\mathbf{x}, \mathbf{y}) = wt(\mathbf{x} + \mathbf{y})$.

Para $\mathbf{x} = (x_1 \ldots x_n)$ e $\mathbf{y} = (y_1 \ldots y_n)$ en F_q^n, sea

$$\mathbf{x} * \mathbf{y} = (x_1 y_1, \ldots, x_n y_n).$$

1.11 Lema. *Si* $\mathbf{x}, \mathbf{y} \in F_2^n$*, entonces*

$$wt(\mathbf{x} + \mathbf{y}) = wt(\mathbf{x}) + wt(\mathbf{y}) - 2wt(\mathbf{x} * \mathbf{y}). \tag{1.3}$$

Demostración. Por (1.2), es suficiente con probar que (1.3) se cumple para $\mathbf{x}, \mathbf{y} \in F_2$. □

El Lema 1.11 implica claramente que $wt(\mathbf{x}) + wt(\mathbf{y}) \geq wt(\mathbf{x} + \mathbf{y})$ para $\mathbf{x}, \mathbf{y} \in F_2^n$, lo cual se cumple también sobre cualquier alfabeto F_q.

Así como está definida la distancia de un código, también existe el concepto de peso de un código.

1.12 Definición. Sea C un código. El *peso mínimo* (*de Hamming*) de C, denotado por $wt(C)$, es el menor de los pesos de las palabras distintas de cero en C.

1.13 Ejemplo. Sea $C = \{001, 011, 111\}$ un código binario.

(i) Las distancias de Hamming para cada pareja de palabras código en C son:

$$d(001, 011) = 1, d(001, 111) = 2 y d(011, 111) = 1.$$

(ii) La distancia y el peso del código C es la mínima de las distancias de Hamming en el mismo, es decir, $d(C) = 1 = wt(C)$.

(iii) Los pesos de las palabras código en C son: $wt(001) = 1$, $wt(011) = 2$ y $wt(111) = 3$.

Supongamos que tenemos un canal de comunicación en el que se transmiten sólo palabras de un código. Pensemos en que hemos recibido la palabra $\mathbf{w}$. Si $\mathbf{w}$ es una palabra código válida, entonces concluiremos que no hubo error alguno en la transmisión, pero si no es así sabremos que ocurrió algún error. En este caso necesitaremos una regla para encontrar la palabra más factible que pudo haber sido enviada. A dicha regla le llamamos *regla de decodificación*. En el capítulo 1 del libro de Ling y Xing (2004) se hace referencia a algunas reglas de decodificación sencillas, pero muy útiles, que se basan en las citadas características de los canales de comunicación: la decodificación por máxima similitud y la decodificación por distancias mínimas.

1.2.1. Decodificación por máxima similitud

Supongamos que se envían las palabras código de un código C por medio de un canal de comunicación. Si la palabra $\mathbf{x}$ es recibida, podemos calcular las probabilidades del canal tales que

$$\mathcal{P}(\mathbf{x} \text{ recibida} | \mathbf{c} \text{ enviada}),$$

para todas las palabras código $\mathbf{c} \in C$. La *regla de decodificación por máxima similitud* (RDMS) concluye que $\mathbf{c_x}$ es la palabra código con mayores probabilidades de haber sido enviada:

$$\mathcal{P}(\mathbf{x} \text{ recibida}|\mathbf{c_x} \text{ enviada}) = \max_{\mathbf{c}\in C} \mathcal{P}(\mathbf{x} \text{ recibida}|\mathbf{c} \text{ enviada}).$$

Existen dos tipos de decodificación por máxima similitud.

(i) *La decodificación completa por máxima similitud*. Si una palabra $\mathbf{x}$ es recibida, aplicar la regla RDMS. Si existe más de una posible palabra enviada, elegirla arbitrariamente.

(ii) *La decodificación incompleta por máxima similitud*. Si una palabra $\mathbf{x}$ es recibida, aplicar la regla RDMS. Si existe más de una posible palabra enviada, solicitar una retransmisión.

1.2.2. Decodificación por distancias mínimas

Supongamos nuevamente que se envían las palabras código de un código C por medio de un canal de comunicación. Si la palabra $\mathbf{x}$ es recibida, la *regla de decodificación por distancias mínimas* (RDDM) decodificará $\mathbf{x}$ como $\mathbf{c_x}$ si $d(\mathbf{x}, \mathbf{c_x})$ es la mínima entre todas las palabras código en C, i.e.,

$$d(\mathbf{x}, \mathbf{c_x}) = \min_{\mathbf{c}\in C} d(\mathbf{x}, \mathbf{c}).$$

Al igual que en el caso de la decodificación por máxima similitud, podemos distinguir las versiones completa e incompleta de este regla. Para una palabra $\mathbf{x}$ recibida, si dos o más palabras $\mathbf{c_x}$ son mínimas e iguales, simplemente se elige a una de ellas de forma arbitraria (RDDM completa), mientras que en la RDDM incompleta en este caso se solicita una retransmisión.

Cabe destacar que cuando $p < 0.5$, las reglas RDMS y RDDM son equivalentes (véase Ling y Xing (2004)), y por esta razón, en adelante nos referiremos únicamente a la regla de decodificación por distancias mínimas.

1.3. Detección y corrección de errores

Conocer la distancia mínima de un código es de suma importancia pues esta métrica minimal nos proporciona, entre otras cosas, una idea de la capacidad del código para detectar y corregir errores, tal como estudiaremos a continuación.

1.14 Definición. Un código de longitud n, tamaño M y distancia d se escribe $C - (n, M, d)$. Los números n, M y d son llamados *parámetros* del código.

1.15 Definición. Sea u un entero positivo. Un código C se dice *detector de u errores* si, siempre que una palabra código incurre en al menos un error y a lo más u errores, la palabra resultante no es una palabra código. Un código C se dice *detector de exactamente u errores* si es detector de u errores, pero no detector de $(u + 1)$ errores.

1.16 Ejemplo. Sea el código binario $C = \{00000, 00111, 11222\}$. Analicemos los cambios necesarios en las coordenadas de cada palabra código de manera que podamos obtener alguna otra palabra código existente en C:

$$\begin{aligned} 00000 &\rightarrow 00111 \text{ necesita cambiar tres bits,} \\ 00000 &\rightarrow 11222 \text{ necesita cambiar cinco bits,} \\ 00111 &\rightarrow 11222 \text{ necesita cambiar cinco bits.} \end{aligned}$$

Por lo tanto, C puede detectar hasta 2 errores. De hecho, C es un código detector de *exactamente* 2 *errores*, pues si cambiamos por 1 las últimas tres coordenadas de 00000 tendremos $00111 \in C$, i.e., C no es un código detector de 3 errores.

El siguiente teorema resulta muy útil al permitir conocer el número de errores que un código es capaz de detectar en función de su distancia.

1.17 Teorema. *Un código C es* detector de u errores *si y sólo si $d(C) \geq u + 1$, es decir, un código con distancia d es un código corrector de exactamente $(d - 1)$ errores.*

Demostración. Supongamos que $d(C) \geq u+1$. Si $\mathbf{c} \in C$ y $\mathbf{x}$ son tales que $1 \leq d(\mathbf{c}, \mathbf{x}) \leq u < d(C)$, entonces $\mathbf{x} \notin C$; por lo tanto, C detecta u errores.

Por otra parte, si $d(C) < u + 1$, i.e., $d(C) \leq u$, entonces existen $\mathbf{c}_1, \mathbf{c}_2 \in C$ tales que $1 \leq d(\mathbf{c}_1, \mathbf{c}_2) = d(C) \leq u$. Por lo que es posible que comencemos con $\mathbf{c}_1$ y se incurra en $d(C)$ errores (donde $1 \leq d(C) \leq u$), de manera que la palabra resultante sea $\mathbf{c}_2$, otra palabra código en C. Por lo tanto, C no es un código detector de u errores. □

1.18 Ejemplo. Retomando el código C del Ejemplo 1.16, observamos que

$$\begin{aligned} d(00000, 00111) &= 3, \\ d(00000, 11222) &= 5, \\ d(00111, 11222) &= 5, \end{aligned}$$

esto es, d(C)=3. Por lo tanto, C detecta exactamente 2 errores, confirmando el resultado obtenido anteriormente.

1.19 Definición. Sea v un entero positivo. Un código C se dice *corrector de v errores* si su distancia mínima permite corregir v o menos errores, asumiendo que sea usada una regla de decodificación incompleta. Se dice que un código C es *corrector de exactamente v errores* si es corrector de v errores, pero no de $(v+1)$ errores.

1.20 Ejemplo. Consideremos el código binario $C = \{000, 111\}$. Supongamos que nos envían una de las dos palabras del código y que usamos la regla de decodificación por distancia mínima. Si ocurrió sólo un error en la transmisión tenemos los siguientes escenarios:

(i) Si enviaron 000, recibimos 100, 010 ó 001. Al emplear la regla de decodificación RDDM en cualquiera de los tres casos, decodificaremos la palabra enviada como 000.

(ii) Si enviaron 111, recibimos 011, 101 ó 110. Al emplear la regla de decodificación RDDM en cualquiera de los tres casos, decodificaremos la palabra enviada como 111.

En cualquiera que sea el caso, es posible corregir un sólo error, por lo que C es corrector de un error.

Si ocurren dos o más errores, la regla RDDM podría fallar. Por ejemplo, si se envía 000 y se recibe 011, entonces 011 será decodificada como 111 con la regla RDMM. Esto es, C es capaz de corregir exactamente un sólo error.

1.21 Teorema. *Un código C es corrector de v errores si y sólo si $d(C) \geq 2v+1$, i.e., un código con distancia d es un código corrector de exactamente $\lfloor (d-1)/2 \rfloor$ errores, donde $\lfloor x \rfloor$ es el mayor entero menor o igual a x.*

Demostración. '$\Leftarrow$' Supongamos que $d(C) \geq 2v+1$. Sean $\mathbf{c}$ la palabra enviada y $\mathbf{x}$ la palabra recibida. Si ocurriesen v o menos errores durante la transmisión, entonces $d(\mathbf{x}, \mathbf{c}) \leq v$. Por (iv) de la Proposición 1.6 (la desigualdad triangular), sabemos que para cualquier $\mathbf{c}' \in C$, en donde $\mathbf{c} \neq \mathbf{c}'$ se cumple que

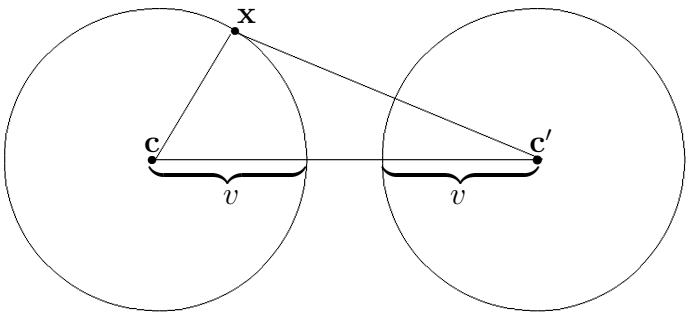
$$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{c}, \mathbf{x}) + d(\mathbf{x}, \mathbf{c}'),$$

véase la Figura 1.3. Por la propiedad (iii) de la misma proposición, sabemos que esto se puede expresar como

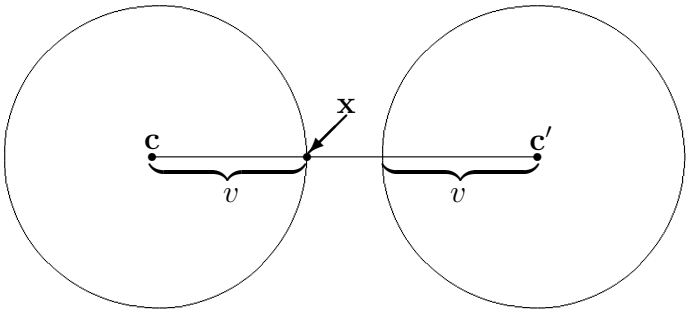
$$d(\mathbf{c}, \mathbf{c}') \leq d(\mathbf{x}, \mathbf{c}) + d(\mathbf{x}, \mathbf{c}'),$$

esto es

$$d(\mathbf{c}, \mathbf{c}') - d(\mathbf{x}, \mathbf{c}) \leq d(\mathbf{x}, \mathbf{c}').$$



a) $d(\mathbf{c}, \mathbf{c}') < d(\mathbf{c}, \mathbf{x}) + d(\mathbf{x}, \mathbf{c}')$



b) $d(\mathbf{c}, \mathbf{c}') = d(\mathbf{c}, \mathbf{x}) + d(\mathbf{x}, \mathbf{c}')$

Fig. 1.3: Interpretaciones de la desigualdad triangular en la corrección de errores

Ahora bien, $2v+1 \leq d(\mathbf{c},\mathbf{c}')$, puesto que tanto $\mathbf{c}$ como $\mathbf{c}'$ son palabras del código y $d(C) \geq 2v+1$. Por lo tanto

$$\begin{aligned} d(\mathbf{c},\mathbf{c}') - d(\mathbf{c},\mathbf{x}) &\leq d(\mathbf{x},\mathbf{c}') \\ 2v+1-v &\leq d(\mathbf{x},\mathbf{c}') \\ v+1 &\leq d(\mathbf{x},\mathbf{c}'). \end{aligned}$$

De donde se concluye que $d(\mathbf{x},\mathbf{c}) \leq v < v+1 \leq d(\mathbf{x},\mathbf{c}')$. Esto significa que al ser recibida $\mathbf{x}$, se decodificará correctamente como $\mathbf{c}$ usando la regla de decodificación RDDM, lo cual prueba que C es corrector de v errores.

'$\Rightarrow$ ' Supongamos que C corrige v errores o menos. Si $d(C) < 2v+1$, entonces existen dos palabras distintas $\mathbf{c}$ y $\mathbf{c}'$ en C tales que $d(\mathbf{c},\mathbf{c}') = d(C) \leq 2v$. Ahora probaremos que, asumiendo que $\mathbf{c}$ sea enviada y a lo más ocurran v errores durante la transmisión, es posible que la regla de decodificación por distancia mínima decodifique la palabra recibida como $\mathbf{c}'$ (otra palabra en el código), o bien, que llegue a un empate (que haya dos distancias mínimas iguales) y que en consecuencia no pueda corregirse ningún error si se emplea una regla de decodificación incompleta (véase Figura 1.4). Esto contradice la suposición de que C es un código corrector de v errores, lo cual prueba que $d(C) \geq 2v+1$.

Si $d(\mathbf{c},\mathbf{c}') < v+1$, entonces $\mathbf{c}$ puede convertirse en $\mathbf{c}'$ si se incurre a lo más en v errores, y esos errores no serán corregidos y ni siquiera serán detectados, puesto que $\mathbf{c}'$ está nuevamente en C. Esto contradice la suposición de que C es corrector de v errores. Por esa razón $d(\mathbf{c},\mathbf{c}') \geq v+1$. Sin pérdida de generalidad, asumiremos que $\mathbf{c}$ y $\mathbf{c}'$ difieren exactamente en las primeras $d = d(C)$ coordenadas, donde $v+1 \leq d \leq 2v$. Si se recibe la palabra

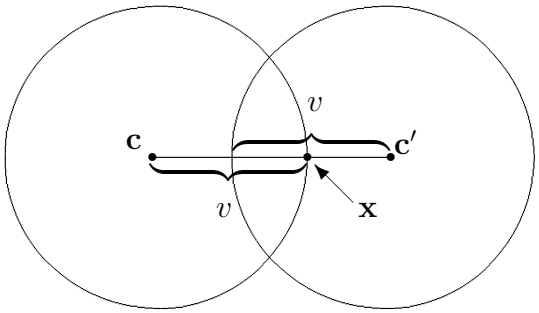
$$\mathbf{x} = \underbrace{x_1 \cdots x_v}_{\text{coincide con } \mathbf{c}'} \; \underbrace{x_{v+1} \cdots x_d}_{\text{coincide con } \mathbf{c}} \; \underbrace{x_{d+1} \cdots x_n}_{\text{coincide con ambos}},$$

entonces tenemos que

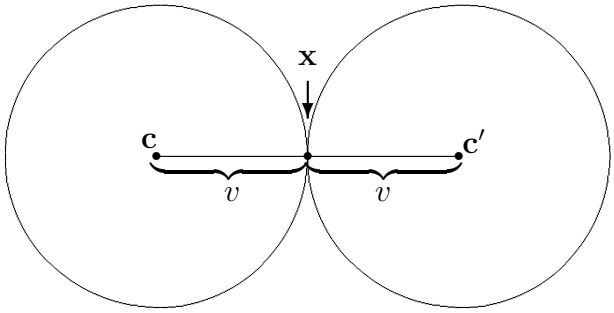
$$d(\mathbf{x},\mathbf{c}') = d - v \leq v = d(\mathbf{x},\mathbf{c}).$$

Se sigue entonces que $d(\mathbf{x},\mathbf{c}') < d(\mathbf{x},\mathbf{c})$, en cuyo caso $\mathbf{x}$ es decodificada incorrectamente como $\mathbf{c}'$, o $d(\mathbf{x},\mathbf{c}) = d(\mathbf{x},\mathbf{c}')$, en cuyo caso se reporta un empate. □

1.22 Ejemplo. Retomando el Ejemplo 1.20, tenemos que $d(000,111) = 3$, i.e., $d(C) = 3$, por lo que el código C es capaz de *corregir* exactamente $\lfloor (3-1)/2 \rfloor = 1$ errores, lo que confirma el resultado anteriormente obtenido.



a) $d(C) < 2v$; $d(\mathbf{c}', \mathbf{x}) < d(\mathbf{c}, \mathbf{x})$
Decodificación incorrecta



b) $d(C) = 2v$; $d(\mathbf{c}', \mathbf{x}) = d(\mathbf{c}, \mathbf{x})$
Empate

Fig. 1.4: Posibles dificultades en la corrección de errores

El libro de Pretzel (1998) expone una estrategia mixta, combinando los últimos dos teoremas, para determinar con facilidad cuántos errores permite detectar un código y al mismo tiempo cuántos permite corregir.

Para concluir esta sección haremos una precisión sobre la probabilidad de que ocurran errores en la transmisión de una palabra código:

(i) Supongamos, sin pérdida de generalidad, que ocurrió un error en las primeras k coordenadas de una palabra código recibida:

$$\mathbf{x} = \underbrace{x_1 \cdots x_k}_{k \text{ con error}} \underbrace{x_{k+1} \cdots x_n}_{n-k \text{ sin error}}.$$

Sea p la probabilidad de que ocurra un error en la transmisión de una coordenada. La probabilidad de que ocurra en la transmisión un error en k coordenadas *fijas* de una palabra de longitud n está dada por

$$p^k(1-p)^{n-k}.$$

(ii) La probabilidad de que ocurra en la transmisión un error en k coordenadas *cualesquiera* de una palabra de longitud n está dada por

$$\binom{n}{k} p^k (1-p)^{n-k}.$$

1.4. Códigos lineales

Los códigos lineales tienen un lugar destacado en la teoría de códigos y algunas razones por las que comúnmente se prefiere usar códigos lineales sobre códigos no lineales son las siguientes:

(i) Puesto que un código lineal es un espacio vectorial, éste puede ser descrito completamente por medio de una base.

(ii) La distancia de un código lineal es igual al menor de los pesos de sus palabras código distintas de cero.

(iii) Los métodos para codificar y decodificar suelen ser más rápidos y sencillos para códigos lineales que para aquéllos no lineales.

1.23 Definición. Sea F_q el campo finito de orden q. Un conjunto no vacío V, junto con las operaciones de adición y multiplicación escalar sobre elementos del campo F_q, es un *espacio vectorial* o *espacio lineal* sobre F_q si satisface todas las siguientes condiciones. Para todas $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in V$ y para todas $\lambda, \mu \in F_q$:

(i) $\mathbf{u} + \mathbf{v} \in V$;

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;

(iii) existe un elemento $\mathbf{0} \in V$ con la propiedad de que $\mathbf{0} + \mathbf{v} = \mathbf{v} = \mathbf{v} + \mathbf{0}$ para toda $\mathbf{v} \in V$;

(iv) para cada $\mathbf{u} \in V$ existe un elemento V, llamado $-\mathbf{u}$, tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0} = (-\mathbf{u}) + \mathbf{u}$;

(v) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

(vi) $\lambda\mathbf{v} \in V$;

(vii) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$, $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u}$;

(viii) $(\lambda\mu)\mathbf{u} = \lambda(\mu\mathbf{u})$;

(ix) si 1 es el neutro multiplicativo en F_q, entonces $1\mathbf{u} = \mathbf{u}$.

1.24 Definición. Sea V un espacio vectorial sobre F_q. Una *combinación lineal* de $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ es un vector de la forma $\lambda_1\mathbf{v}_1 \ldots \lambda_r\mathbf{v}_r$, donde $\lambda_1, \ldots, \lambda_r \in F_q$ son escalares.

1.25 Definición. Sea V un espacio vectorial sobre F_q. El conjunto de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ en V se dice que es *linealmente independiente* si

$$\lambda_1\mathbf{v}_1 + \ldots + \lambda_r\mathbf{v}_r = \mathbf{0} \Rightarrow \lambda_1 = \ldots = \lambda_r = 0.$$

El conjunto es *linealmente dependiente* si no es linealmente independiente; i.e., si hay $\lambda_1, \ldots, \lambda_r \in F_q$, no todos cero, tales que $\lambda_1\mathbf{v}_1 + \ldots + \lambda_r\mathbf{v}_r = \mathbf{0}$.

1.26 Definición. Sea V un espacio vectorial sobre F_q y sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un subconjunto no vacío de V. La *expansión lineal* de S se define como

$$< S >= \{\lambda_1\mathbf{v}_1 + \ldots + \lambda_k\mathbf{v}_k : \lambda_i \in F_q\}.$$

Si $S = \emptyset$ definimos $< S >= \{\mathbf{0}\}$. Es sencillo verificar que $< S >$ es un subespacio de V, llamado el *subespacio generado* por S. Dado un subespacio C de V, el subconjunto S de C se llama *conjunto generador* de C si $C =< S >$.

1.27 Definición. Sea V un espacio vectorial sobre F_q. Un subconjunto no vacío $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ de V se llama *base* de V si $V =< B >$ y B es linealmente independiente.

Es importante destacar que si $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base de V, entonces cualquier vector $\mathbf{v} \in V$ puede expresarse como una combinación lineal única de vectores en B, esto es, existen $\lambda_1, \lambda_1, \ldots, \lambda_k \in F_q$ únicas tales que

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_1\mathbf{v}_k.$$

Por otra parte, un espacio vectorial V sobre un campo finito F_q puede tener múltiples bases, pero todas esas bases constan del mismo múmero de elementos. Este número de elementos es llamado *dimensión* de V sobre F_q, y frecuentemente se denota como $dim(V)$.

1.28 Definición. Sea $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, $\mathbf{w} = (w_1, w_2, \ldots, w_n) \in F_q^n$.

(i) El *producto escalar* o *producto euclidiano interno* de $\mathbf{v}$ y $\mathbf{w}$ se define como

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \ldots + v_nw_n \in F_q.$$

(ii) Se dice que dos vectores $\mathbf{v}$ y $\mathbf{w}$ son ortogonales si $\mathbf{v} \cdot \mathbf{w} = 0$.

(iii) Sea S un subconjunto no vacío de F_q^n. El *complemento ortogonal* $S^\perp$ de S se define como

$$S^\perp = \{\mathbf{v} \in F_q^n : \mathbf{v} \cdot \mathbf{s} = 0 \text{ para toda } \mathbf{s} \in S\}.$$

Si $S = \emptyset$, entonces definimos $S^\perp = F_q^n$.

1.29 Definición. *Sea S un subconjunto de F_q^n, entonces tenemos*

$$dim(< S >) + dim(S^\perp) = n.$$

Demostración. El caso trivial es obvio pues $< S >= \{\mathbf{0}\}$.

Sea $dim(< S >) = k \geq 1$ y supongamos que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es una base de $< S >$. Necesitamos probar que $dim(S^\perp) = dim(< S >^\perp) = n - k$.

Sabemos que $\mathbf{x} \in S^\perp$ si y sólo si

$$\mathbf{v}_1 \cdot \mathbf{x} = \ldots = \mathbf{v}_k \cdot \mathbf{x} = 0,$$

lo que equivale a decir que $\mathbf{x}$ satisface $A\mathbf{x}^T = \mathbf{0}$, donde A es la matriz de $k \times n$ cuyo i-ésimo renglón es $\mathbf{v}_i$. Los renglones de A son linealmente independientes, de forma que $A\mathbf{x}^T = \mathbf{0}$, es un sistema lineal de k ecuaciones linealmente independientes de n variables. Del álgebra lineal, se sabe que dicho sistema admite un espacio de solución de $n - k$. □

1.30 Definición. Un *código lineal* C de longitud n sobre F_q es un subespacio de F_q^n.

1.31 Definición. Sea C un código lineal en F_q^n

(i) El *código dual* de C es $C^\perp$, el complemento ortogonal del subespacio C de F_q^n.

(ii) La *dimensión* del código lineal C es la dimensión de C como un espacio vectorial sobre F_q, i.e., $dim(C)$.

1.32 Teorema. *Sea C un código lineal de longitud n sobre F_q. Entonces,*

(i) $|C| = q^{dim(C)}$, *i.e.,* $dim(C) = \log_q |C|$;

(ii) $C^\perp$ *es un código lineal y* $dim(C) + dim(C^\perp) = n$;

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

Demostración. (i) Si $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$ es una base de C, entonces

$$C = \{\lambda_1 \mathbf{c}_1 + \ldots + \lambda_k \mathbf{c}_k : \lambda_1, \ldots, \lambda_k \in F_q\}.$$

Puesto que $|F_q| = q$, hay exactamente q posibilidades para cada $\lambda_1, \ldots, \lambda_k$; por lo que C tiene exactamente $q^k = q^{dim(C)}$ elementos.

(ii) Es sencillo verificar que $S^\perp$ es siempre un subespacio del espacio vectorial F_q^n para todo subconjunto S de F_q^n, y que $< S >^\perp = S^\perp$. La prueba se sigue de emplear el Teorema 1.29 haciendo $C = S$.

(iii) Utilizando la igualdad en (ii) y una igualdad similar reemplazando C por $C^\perp$, obtenemos $dim(C) = dim((C^\perp)^\perp)$. Para probar (iii), basta mostrar que $C \subseteq (C^\perp)^\perp$.

Sea $\mathbf{c} \in C$. Para mostrar que $\mathbf{c} \in (C^\perp)^\perp$, debemos probar que $\mathbf{c} \cdot \mathbf{x} = 0$ para toda $\mathbf{x} \in C^\perp$. Puesto que $\mathbf{c} \in C$ y $\mathbf{x} \in C^\perp$, por definición de $C^\perp$, se sigue que $\mathbf{c} \cdot \mathbf{x} = 0$. Por lo que (iii) queda probado.

□

Un código lineal C de longitud n y de dimensión k sobre F_q es frecuentemente llamado código q-ario $[n, k]$ o, si q es claro en el contexto, simplemente código $[n, k]$. También se denota como código lineal (n, q^k). Si se conoce la distancia d de un código C también se utiliza la notación $[n, k, d]$ para especificar los parámetros del código lineal.

1.33 Definición. Sea C un código lineal.

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

(ii) C es *auto dual* si $C = C^{\perp}$.

1.34 Proposición. *La dimensión de un código auto ortogonal de longitud n debe ser $\leq n/2$, y la dimensión de un código auto dual de longitud n es $n/2$.*

Demostración. Esta proposición es consecuencia inmediata del Teorema 1.32 (ii) y de las definiciones de códigos auto ortogonales y auto duales. □

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

Demostración. Por definición, existen $\mathbf{x}', \mathbf{y}' \in C$ tales que $d(\mathbf{x}', \mathbf{y}') = d(C)$, y por el Lema 1.9 se sigue que

$$d(C) = d(\mathbf{x}', \mathbf{y}') = wt(\mathbf{x}' - \mathbf{y}') \geq wt(C),$$

puesto que $\mathbf{x}' - \mathbf{y}' \in C$.

Por otra parte, hay una $\mathbf{z} \in C \setminus \{\mathbf{0}\}$ tal que $wt(C) = wt(\mathbf{z})$, por lo que

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

completando así la prueba. □

1.4.1. Bases de códigos lineales y algoritmos

Puesto que un código lineal es un espacio vectorial, todos sus elementos pueden ser descritos en términos de una base. En esta sección abordaremos algunos algoritmos que proporcionan una base ya sea para un código lineal o bien para su dual.

Primero retomaremos algunos aspectos fundamentales del álgebra lineal:

1.36 Definición. Sea A una matriz sobre F_q; se considera una *operación elemental de renglón* realizada sobre A cualquiera de estas operaciones:

(i) intercambiar dos columnas,

(ii) multiplicar un renglón por un escalar distinto de cero,

(iii) reemplazar un renglón por su suma con otro renglón previamente multiplicado por un escalar.

1.37 Definición. Dos matrices son *equivalentes por renglón* si una puede ser obtenida a partir de la otra partiendo de una secuencia de operaciones elementales.

Los siguientes son hechos bien conocidos del álgebra lineal:

(i) Toda matriz M sobre F_q puede escribirse en la *forma escalonada de renglón* (FER) o en la *forma escalonada reducida de renglón* (FERR) por medio de una secuencia de operaciones elementales de renglón. En otras palabras, una matriz es equivalente por renglón a otra en la forma FER o en forma FERR.

(ii) Para una matriz dada su FERR es única, pero puede tener diferentes FERs.

A continuación se explican tres algoritmos para obtener la base de un código lineal.

Algoritmo 1.1

Entrada: Un subconjunto no vacío de F_q^n.

Salida: Una base para $C =< S >$, el código lineal generado por S.

Descripción: Formar la matriz A cuyos renglones sean las palabras en S. Por medio de operaciones de renglón elementales, encontrar la FER de A. Los renglones distintos de cero de la FER de A forman una base de C.

Algoritmo 1.2

Entrada: Un subconjunto no vacío de F_q^n.

Salida: Una base para $C =< S >$, el código lineal generado por S.

Descripción: Formar la matriz A cuyos renglones sean las palabras en S. Usar operaciones de renglón elementales para encontrar la FERR de A y localizar las columnas líderes en la FER (aquellas que tienen un elemento igual con 1, y cuyos elementos son todos iguales con 0 por debajo del 1). Las columnas originales de A correspondientes a las columnas líderes forman una base para C.

La base que se obtiene con el Algoritmo 1.2. proporciona un subconjunto del conjunto dado S, lo cual no ocurre necesariamente en el Algoritmo 1.1.

Algoritmo 1.3

Entrada: Un subconjunto no vacío de F_q^n.

Salida: Una base para $C^{\perp}$, donde $C =< S >$, el código lineal generado por S.

Descripción: Formar la matriz A cuyos renglones sean las palabras en S. Utilizar operaciones de renglón elementales para encontrar la FERR de la matriz A. Sea G la matriz de $k \times n$ que consta de todos los renglones distintos de cero de A:

$$A \rightarrow \begin{pmatrix} G \\ O \end{pmatrix}$$

donde O denota la matriz cero.

La matriz G contiene k columnas líderes. Permutar las columnas de G para formar

$$G' = (I_k|X),$$

donde I_k denota la matriz identidad de $k \times k$. Formar la matriz H' como sigue:

$$H' = (-X^T|I_{n-k}),$$

donde X^T denota la transpuesta de X.

Aplicar la inversa de la permutación aplicada a las columnas de G a las columnas de H' para formar H. Entonces, los renglones de H forman una base de $C^{\perp}$.

Note que el Algoritmo 1.3. proporciona también una base para C puesto que incluye al Algoritmo 1.1. Los principios que fundamentan el Algoritmo 1.3. se explican a continuación.

1.4.2. Matrices generadora y parity check

Si se conoce la base de un código lineal, podemos describir las palabras código del mismo de forma explícita. La teoría de códigos representa con frecuencia la base de un código lineal por medio de una matriz llamada *matriz generadora*, y la base de su código dual por medio de una matriz llamada *parity check*. Estas dos matrices juegan un papel muy importante en la teoría de códigos.

1.38 Definición. (i) Una *matriz generadora* para un código lineal C es una matriz G cuyos renglones forman una base para C.

(ii) Una *matriz parity check* H para un código lineal C es la matriz generadora del código dual $C^{\perp}$.

1.39 Ejemplo. Encontremos una matriz generadora estándar G' para el código lineal binario C' equivalente al código lineal C con matriz parity check

$$H = \begin{array}{c} \begin{matrix} a & b & c & d & e \end{matrix} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}.$$

Primero, encontremos la forma estándar de H, esto es, $H' = (-X^T|I_{n-k})$ cambiando el orden de las columnas de H:

$$H' = \begin{array}{c} \begin{matrix} b & e & a & c & d \end{matrix} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

de donde se aprecia que

$$-X^T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = X^T$$

puesto que el código es binario.

Entonces

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Ahora, podemos encontrar la forma estándar de la matriz generadora del código C', $G' = (I_k|X)$:

$$G' = \begin{array}{c} \begin{matrix} b & e & a & c & d \end{matrix} \\ \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \end{array}.$$

De hecho, podemos encontrar a partir de esta información la matriz generadora G del código C, ordenando las columnas que acomodamos en un principio:

$$G = \begin{array}{c} \begin{matrix} a & b & c & d & e \end{matrix} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array}.$$

Dicho sea de paso, se observa que:

(i) La matriz H tiene dimensión 3, puesto que H cuenta con tres renglones, por lo que $dim(C^{\perp}) = 3$.

(ii) Haciendo una revisión exhaustiva, la suma de cualesquiera dos columnas de H es $\mathbf{x}^T \neq \mathbf{0}$. Sin embargo, si se suman las columnas a, b y c se obtiene el vector $\mathbf{0}$. Esto quiere decir que la matriz H tiene tres columnas linealmente dependientes, por lo que $d(C) = 3$.

(iii) De la matriz generadora de C, se observa que $dim(C) = 2 = k$, por lo que el código binario C cuenta con $q^k = 2^2 = 4$ elementos.

(iv) La longitud de las palabras código en C es $n = 5$.

Por lo que, por una parte, puede verificarse que

$$dim(C) + dim(C^{\perp}) = n,$$

tal como comentamos con anterioridad. Y por otra parte, concluimos que los parámetros del código son $C - [5, 2, 3]$, de acuerdo con la notación que hemos venido manejando.

Cabe destacar lo siguiente:

(i) Si C es un código lineal con parámetros $[n, k]$, entonces su matriz generadora es una matriz de $k \times n$ y su matriz parity check es una matriz de $(n - k) \times n$.

(ii) El Algoritmo 1.3 permite encontrar las matrices generadora y parity check de un código lineal.

(iii) Puesto que el número de bases para un espacio vectorial usualmente es mayor a uno, la matriz generadora de un código lineal no es única y, en consecuencia, tampoco lo es la matriz generadora de su código dual. Además, aun considerando una matriz generadora fija, al permutar sus renglones se obtendría una matriz generadora diferente igualmente válida para el código lineal.

(iv) Los renglones de la matriz generadora son linealmente independientes; asimismo, los renglones de la matriz parity check. Para probar que la matriz G de $k \times n$ es en efecto una matriz generadora para el código lineal C con parámetros $[n, k]$, basta mostrar que los renglones de G son palabras código de C y que son linealmente independientes.

1.40 Definición. (i) Se dice que una matriz generadora en la forma $(I_k|X)$ se encuentra en su *forma estándar*.

(ii) Se dice que una matriz parity check en la forma $(Y|I_{n-k})$ se encuentra en su *forma estándar*.

Algunas ventajas de encontrar una forma estándar para la matriz generadora G son las siguientes:

(i) Un código lineal C tiene matriz generadora G en su forma estándar, $G = (I|X)$, entonces el Algoritmo 1.3. a la vez proporciona

$$H = \left(-X^T|I\right)$$

como matriz parity check para C.

(ii) Si un código lineal C con parámetros $[n, k, d]$ tiene matriz generadora G en su forma estándar, $G = (I|X)$, entonces el recuperar el mensaje $\mathbf{u}$ a partir de la palabra código $\mathbf{v} = \mathbf{u}G$ es trivial, puesto que

$$\mathbf{v} = \mathbf{u}G = \mathbf{u}(I|X) = (\mathbf{u}, \mathbf{u}X);$$

es decir, los primeros k dígitos de la palabra código $\mathbf{v} = \mathbf{u}G$ son los que nos dan la información suficiente para conocer el mensaje $\mathbf{u}$. Los $n - k$ dígitos restantes son comúnmente llamados *dígitos verificadores*, mismos que representan la *redundancia* que ha sido agregada al mensaje a manera de protección contra el ruido.

1.41 Lema. *Sea C un código lineal sobre F_q, con matriz generadora G. Entonces $\mathbf{v} \in F_q^n$ pertenece a $C^\perp$ si y sólo si $\mathbf{v}$ es ortogonal a todo renglón de G, i.e., $\mathbf{v} \in C^\perp \Leftrightarrow \mathbf{v}G^T = \mathbf{0}$. En particular, dada una matriz H de $(n-k) \times n$, entonces H es una matriz parity check para C si y sólo si los renglones de H son linealmente independientes y $HG^t = 0$*

Demostración. Sea $\mathbf{r}_i$ el i-ésimo renglón de G. En particular, $\mathbf{r}_i \in C$ para toda $1 \leq i \leq k$, y toda $\mathbf{c} \in C$ puede ser escrita como

$$\mathbf{c} = \lambda_1\mathbf{r}_1 + \ldots + \lambda_k\mathbf{r}_k,$$

donde $\lambda_1, \ldots, \lambda_k \in F_q$. Si $\mathbf{v} \in C^\perp$, entonces $\mathbf{v} \cdot \mathbf{c} = 0$ para toda $\mathbf{c} \in C$. En particular, $\mathbf{v}$ es ortogonal a $\mathbf{r}_i$ para toda $1 \leq i \leq k$; i.e., $\mathbf{v}G^T = \mathbf{0}$.

A la inversa, si $\mathbf{v} \cdot \mathbf{r_i} = 0$ para toda $1 \leq i \leq k$, entonces para cualquier $\mathbf{c} = \lambda_1\mathbf{r}_1 + \ldots + \lambda_k\mathbf{r}_k \in C$,

$$\mathbf{v} \cdot \mathbf{c} = \lambda_1(\mathbf{v} \cdot \mathbf{r}_1) + \ldots + \lambda_k(\mathbf{v} \cdot \mathbf{r}_k) = 0.$$

La última parte del lema se fundamenta en lo siguiente. Si H es una matriz parity check para el código C, entonces los renglones de H son linealmente independientes por definición. Puesto que las columnas de H son palabras código de $C^\perp$, se sigue de

la declaración anterior que $HG^T = 0$. Asimismo, si $HG^T = 0$, entonces la declaración anterior muestra que los renglones de H, y por tanto el espacio de renglones de H, están contenidos en $C^\perp$. Puesto que los renglones de H son linealmente independientes, el espacio de renglones de H tiene dimensión $n - k$, por lo que el espacio de renglones de H es en realidad $C^\perp$. En otras palabras, H es una matriz parity check para C. □

Otra forma equivalente de plantear el lema anterior es la siguiente:

Sea C un código lineal con parámetros $[n, k]$ sobre F_q, con matriz parity check H. Entonces $\mathbf{v} \in F_q^n$ pertenece a C si y sólo si $\mathbf{v}$ es ortogonal a todo renglón de H; i.e., $\mathbf{v} \in C \Leftrightarrow \mathbf{v}H^T = \mathbf{0}$. En particular, dada una matriz G de $k \times n$, entonces G es una matriz generadora para C si y sólo si los renglones de G son linealmente independientes y $GH^T = 0$.

Una de las consecuencias del lema anterior es el siguiente teorema que relaciona la distancia de un código lineal C con las propiedades de una matriz parity check de C.

1.42 Teorema. *Sea C un código lineal y sea H una matriz parity check para C. Entonces*

(i) C tiene distancia $\geq d$ si y sólo si cualesquiera $d-1$ columnas de H son linealmente independientes; y

(ii) C tiene distancia $\leq d$ si y sólo si H tiene d columnas que son linealmente independientes.

Demostración. Sea $\mathbf{v} = (v_1, \ldots, v_n) \in C$ una palabra de peso $e \geq 0$. Supongamos que las coordenadas de $\mathbf{v}$ distintas de cero se encuentran en las posiciones $i_1, \ldots, i_e$ de manera que $v_j = 0$ si $j \notin \{i_1, \ldots, i_e\}$. Sea $\mathbf{c}_i$ $(1 \leq i \leq n)$ la i-ésima columna de H.

Por el Lema 1.41, o mejor dicho, por el replanteamiento que hicimos posteriormente al mismo, C contiene la palabra $\mathbf{v} = (v_1, \ldots, v_n)$ de peso e (cuyas coordenadas distintas de cero son $v_{i_1}, \ldots, v_{i_e}$) si y sólo si

$$\mathbf{0} = \mathbf{v}H^T = v_{i_1}\mathbf{c}_{i_1}^T + \ldots + v_{i_e}\mathbf{c}_{i_e}^T,$$

lo cual es cierto si y sólo si existen e columnas de H (concretamente, $\mathbf{c}_{i_1}, \ldots, \mathbf{c}_{i_e}$) que son linealmente dependientes.

Decir que la distancia de C es $\geq d$ equivale a decir que C no contiene ninguna palabra distinta de cero de peso $\leq d-1$, i.e., cualesquiera $\leq d-1$ columnas de H son linealmente independientes. Esto prueba la afirmación (i).

De manera semejante, decir que la distancia de C es $\leq d$ equivale a decir que H tiene $\leq d$ columnas (y por lo tanto d columnas) que son linealmente dependientes. Esto prueba la afirmación (ii). □

Cuando d es pequeña, el siguiente corolario puede ser de utilidad para calcular su valor. Este corolario se despende del Teorema 1.42.

1.43 Corolario. *Sea C un código lineal y sea H una matriz parity check para C. Las siguientes afirmaciones son equivalentes:*

(i) C tiene distancia d;

(ii) cualesquiera $d-1$ columnas de H son linealmente independientes y H tiene d columnas que son linealmente dependientes.

1.44 Teorema. *Si $G = (I_k|X)$ es la forma estándar de la matriz generadora de un código $C - [n, k]$, entonces la matriz parity check de C es $H = (-X^T|I_{n-k})$.*

Demostración. La ecuación $HG^T = O$ se satisface. Considerando las últimas $n-k$ coordenadas, los renglones de H son linealmente independientes. Por lo tanto, la conclusión se sigue del Lema 1.41. □

El Teorema 1.44 prueba que el Algoritmo 1.3 proporciona una base para el código dual de C, en otras palabras, proporciona la matriz parity check de C.

1.4.3. Equivalencia de códigos lineales

A pesar de que existen ciertos códigos lineales que pueden no tener una matriz generadora en su forma estándar, después de algunas permutaciones de las coordenadas de las palabras código que la componen y, posiblemente, multiplicando ciertas coordenadas por escalares distintos de cero, podemos encontrar un nuevo código que sí tenga una matriz generadora.

1.45 Definición. Dos códigos de parámetros (n, M) sobre F_q son *equivalentes* si puede obtenerse uno a partir del otro por medio de una combinación de operaciones del siguiente tipo:

(i) permutación de los n dígitos de las palabras código;

(ii) multiplicación de los símbolos que aparecen en una posición fija por un escalar distinto de cero.

1.46 Teorema. *Cualquier código lineal C es equivalente a un código lineal C' con una matriz generadora en su forma estándar.*

Demostración. Si G es una matriz generadora para C, pongamos G en su FERR. Acomodemos las columnas de la FERR de manera que las columnas líderes formen una matriz identidad. Como resultado se tendrá la matriz G', en su forma estándar que es la matriz generadora de un código C' equivalente al código C. □

El Teorema 1.46 es en realidad la primera parte del Algoritmo 1.3.

1.4.4. Codificación de un código lineal

Sea C un código lineal con parámetros $[n, k, d]$ sobre el campo finito F_q. Cada palabra código de C puede representar una pieza de información, por lo que C puede representar q^k pedazos de información distintos. Una vez que se fija una base $\mathbf{r}_1 \ldots, \mathbf{r}_k$ para C, cada palabra código $\mathbf{v}$ (cada pieza de información) puede ser escrita de forma única por medio de una combinación lineal,

$$\mathbf{v} = u_1\mathbf{r}_1 + \ldots + u_k\mathbf{r}_k,$$

donde $u_1, \ldots, u_k \in F_q$.

De manera equivalente, podemos establecer un conjunto G como matriz generadora de C cuyo i-ésimo renglón es el vector $\mathbf{r}_i$ de la base elegida. Dado un vector $\mathbf{u} = (u_1, \ldots, u_k) \in F_q^k$, entonces

$$\mathbf{v} = \mathbf{u}G = u_1\mathbf{r}_1 + \ldots + u_k\mathbf{r}_k$$

es una palabra código en C. Por otra parte, cualquier $\mathbf{v} \in C$ puede escribirse de forma única como $\mathbf{v} = \mathbf{u}G$, donde $\mathbf{u} = (u_1, \ldots, u_k) \in F_q^k$. Por lo tanto, toda palabra código $\mathbf{u} \in F_q^k$ puede codificarse como $\mathbf{v} = \mathbf{u}G$.

El proceso de representar los elementos $\mathbf{u}$ en F_q^k como palabras código $\mathbf{v} = \mathbf{u}G$ en C se conoce como *codificar*.

1.47 Ejemplo. Asignemos las siguientes letras a las palabras en F_2^3 como sigue:

000	001	010	011	100	101	110	111
A	E	I	O	N	S	T	P

Sea C el código lineal binario con matriz generadora

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

Usaremos G para codificar el mensaje EINSTEIN. Por una parte, sabemos que las palabras en F_2^3 asociadas directamente al mensaje EINSTEIN son

001	010	100	101	110	001	010	100
E	I	N	S	T	E	I	N

Además, sabemos que la codificación, y dicho sea de paso, la decodificación de un mensaje es más sencilla utilizando una matriz generadora en su forma estándar. La forma estándar de G, en este caso, es

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

Ahora, haremos $\mathbf{v} = \mathbf{u}G'$, para cada una de las u_i de nuestro alfabeto:

$$v_1 = u_1G' = (001)\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (00101)$$

$$v_2 = u_2G' = (010)\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (01001)$$

$$v_3 = u_3G' = (100)\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (10011)$$

$$v_4 = u_4G' = (101)\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (10110)$$

y

$$v_5 = u_5G' = (110)\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = (11010).$$

Note que la simplicidad de usar la matriz G' radica en que la palabra codificada queda igual a la inicial, salvo porque le son añadidos dos dígitos, los dígitos de redundancia, de forma tal que es sumamente fácil también decodificarla.

Por lo que el mensaje EINSTEIN,empleando la matriz G' especificada, queda codificado como

$$00101 \quad 01001 \quad 10011 \quad 10110 \quad 11010 \quad 00101 \quad 01001 \quad 10011$$

1.4.5. Decodificación de un código lineal

Un código tiene uso práctico si se le aplica un esquema de decodificación eficiente. A continuación explicaremos brevemente algunas decodificaciones sencillas, pero elegantes y eficientes.

Decodificación por el vecino más cercano

Comenzaremos por establecer una definición para el concepto de co-conjunto.

1.48 Definición. Sea C un código lineal de longitud n sobre F_q, y sea $\mathbf{u} \in F_q^n$ cualquier vector de longitud n; definimos el *co-conjunto* de C determinado por $\mathbf{u}$ como el conjunto

$$C + \mathbf{u} = \{\mathbf{v} + \mathbf{u} : \mathbf{v} \in C\} = \mathbf{u} + C.$$

1.49 Teorema. *Sea C un código lineal con parámetros $[n, k, d]$ sobre el campo finito F_q. Entonces:*

(i) todo vector de F_q^n está contenido en algún co-conjunto de C;

(ii) para todo $\mathbf{u} \in F_q^n$, $|C + \mathbf{u}| = |C| = q^k$;

(iii) para todos $\mathbf{u}, \mathbf{v} \in F_q^n$, $\mathbf{u} \in C + \mathbf{v}$ implica que $c + \mathbf{u} = C + \mathbf{v}$;

(iv) dos co-conjuntos son idénticos, o bien, tienen intersección nula;

(v) hay q^{n-k} co-conjuntos diferentes de C;

(vi) para todos $\mathbf{u}, \mathbf{v} \in F_q^n$, $\mathbf{u} - \mathbf{v} \in C$ si y sólo si $\mathbf{u}$ y $\mathbf{v}$ están en el mismo co-conjunto.

Demostración. (i) El vector $\mathbf{v} \in F_q^n$ está contenido en el co-conjunto $C + \mathbf{v}$.

(ii) Por definición, $C + \mathbf{u}$ tiene a lo más $|C| = q^k$ elementos. Dos elementos $\mathbf{c} + \mathbf{u}$ y $\mathbf{c}' + \mathbf{u}$ de $C + \mathbf{u}$ son iguales si y sólo si $\mathbf{c} = \mathbf{c}'$, por lo tanto $|C + \mathbf{u}| = |C| = q^k$.

(iii) Se sigue de la definición de $C + \mathbf{v}$ que $C + \mathbf{u} \subseteq C + \mathbf{v}$. Entonces, por (ii), $C + \mathbf{u} = C + \mathbf{v}$.

(iv) Consideremos dos co-conjuntos $C + \mathbf{u}$ y $C + \mathbf{v}$, y supongamos que $\mathbf{x} \in (C + \mathbf{u}) \cap (C + \mathbf{v})$. Puesto que $\mathbf{x} \in C + \mathbf{u}$, (iii) muestra que $C + \mathbf{u} = C + \mathbf{x}$. De manera similar, puesto que $\mathbf{x} \in C + \mathbf{v}$, se sigue que $C + \mathbf{v} = C + \mathbf{x}$. Por lo tanto, $C + \mathbf{u} = C + \mathbf{v}$.

(v) Se demuestra directamente de (i), (ii) y (iv).

(vi) Si $\mathbf{u} - \mathbf{v} = \mathbf{c} \in C$, entonces $\mathbf{u} = \mathbf{c} + \mathbf{v} \in C + \mathbf{v}$, por lo que $C + \mathbf{u} = C + \mathbf{v}$. Por la prueba de (i), $\mathbf{u} \in C + \mathbf{u}$ y $\mathbf{v} \in C + \mathbf{v}$, por lo que $\mathbf{u}$ y $\mathbf{v}$ están en el mismo co-conjunto.

Por otra parte, supongamos que $\mathbf{u}$, $\mathbf{v}$ están ambos en el co-conjunto $C + \mathbf{x}$. Entonces $\mathbf{u} = \mathbf{c} + \mathbf{x}$ y $\mathbf{v} = \mathbf{c}' + \mathbf{x}$, para algunos $\mathbf{c}, \mathbf{c}' \in C$. Por lo tanto, $\mathbf{u} - \mathbf{v} = \mathbf{c} - \mathbf{c}' \in C$.

□

1.50 Definición. A la palabra con el menor peso (de Hamming) en un co-conjunto se le llama *co-conjunto líder*.

1.51 Ejemplo. Busquemos cuántos co-conjuntos pueden obtenerse para el código lineal binario $C = \{00000, 10001, 01010, 11011, 00100, 10101, 01110, 11111\}$, así como los elementos de los que consta cada uno de ellos.

Es claro que uno de los co-conjuntos de C es el propio C ya que $00000 + C = C$. Comenzamos sumando a cada palabra del código C una palabra de peso pequeño:

$$00001 + C = \{00001, 10000, 01011, 11010, 00101, 10100, 01111, 11110\},$$
$$00010 + C = \{00010, 10011, 01000, 11001, 00110, 10111, 01100, 11101\}.$$

Puesto que en los co-conjuntos previos ya se encuentran todas las palabras de peso 1, encontramos otro co-conjunto sumando a C cualquier palabra de peso 2 que no se encuentre en los co-conjuntos anteriores, por ejemplo 11000:

$$11000 + C = \{11000, 01001, 10010, 00011, 11100, 01101, 10110, 00111\}.$$

Esto es C tiene 4 co-conjuntos con 8 elementos cada uno, para acumular un total de $2^5 = 32$ arreglos de 5 bits en $q = 2$.

Sea C un código lineal. Asumamos que se transmite la palabra código $\mathbf{v}$ y que se recibe la palabra $\mathbf{w}$, de forma tal que hay un *patrón de error* de este tipo:

$$\mathbf{e} = \mathbf{w} - \mathbf{v} \in \mathbf{w} + C.$$

Entonces, $\mathbf{w} - \mathbf{e} = \mathbf{v} \in C$, por lo que, empleando (vi) del Teorema 1.49, el patrón de error $\mathbf{e}$ y la palabra recibida $\mathbf{w}$ pertenecen al mismo co-conjunto.

Puesto que los patrones de error de peso pequeño tienen mayor probabilidad de ocurrencia, la decodificación por medio del vecino más cercano funciona para la decodificación de códigos lineales de la siguiente manera: una vez que es recibida la palabra $\mathbf{w}$, elegimos la palabra $\mathbf{e}$ de menor peso en el co-conjunto $\mathbf{w} + C$ y concluimos que $\mathbf{v} = \mathbf{w} - \mathbf{e}$ es la palabra que fue transmitida originalmente. [4]

Decodificación por síndromes

El esquema de decodificación basado en un arreglo estándar trabaja de forma razonable cuando la longitud n de un código lineal es pequeña, pero puede tomar una cantidad de tiempo considerable conforme n se hace grande. Puede ahorrarse tiempo haciendo uso de un síndrome para identificar el co-conjunto al que pertenece la palabra recibida.

1.52 Definición. Sea C un código lineal con parámetros $[n, k, d]$ sobre F_q y sea H la matriz parity check para C. Para cualquier $\mathbf{w} \in F_q^n$, el *síndrome* de $\mathbf{w}$ es la palabra $S(\mathbf{w}) = \mathbf{w}H^T \in F_q^{n-k}$, esto es, el síndrome de $\mathbf{w}$ depende de la elección de la matriz parity check para C.

1.53 Teorema. *Sea C un código lineal con parámetros $[n, k, d]$ y sea H una matriz parity check para C. Para $\mathbf{u}$, $\mathbf{v} \in F_q^n$, tenemos que*

(i) $S(\mathbf{u} + \mathbf{v}) = S(\mathbf{u}) + S(\mathbf{v})$;

(ii) $S(\mathbf{u}) = \mathbf{0}$ *si y sólo si* $\mathbf{u}$ *es una palabra código en* C;

(iii) $S(\mathbf{u}) = S(\mathbf{v})$ *si y sólo si* $\mathbf{u}$ *y* $\mathbf{v}$ *pertenecen al mismo co-conjunto de* C.

Demostración. (i) es una consecuencia inmediata de la definición de síndrome.

(ii) Por la definición de síndrome, $S(\mathbf{u}) = \mathbf{0}$ si y sólo si $\mathbf{u}H^T = \mathbf{0}$, lo cual, por el replanteamiento del Lema 1.41, equivale a que $\mathbf{u} \in C$.

[4] Recordemos que se considera que el canal de comunicación es eficiente en la medida en la que la ocurrencia de errores en la transmisión es considerablemente pequeña.

(iii) se sigue de (i), (ii) y el Teorema 1.49 (vi).

□

El Teorema 1.53 establece que podemos identificar un co-conjunto por medio de su síndrome; a la inversa, todas las palabras código en un determinado co-conjunto tienen el mismo síndrome, esto es, el síndrome de un co-conjunto es el síndrome de todas las palabras código en dicho co-conjunto. Esto quiere decir que existe una correspondencia uno a uno entre los co-conjuntos y los síndromes.

Puesto que los síndromes se encuentran en F_q^{n-k}, existen a lo más q^{n-k} síndromes. El Teorema 1.49 (v) dice que hay q^{n-k} co-conjuntos, por lo que hay q^{n-k} síndromes, todos distintos entre sí. Por lo tanto, todos los vectores en F_q^{n-k} son síndromes.

1.54 Definición. La tabla que asocia cada co-conjunto líder con su síndrome se llama *tabla de síndromes* o *arreglo de decodificación estándar*.

A continuación se presentan los pasos para construir una tabla de síndromes bajo el esquema de la decodificación *completa* por medio del vecino más cercano.

Pasos para construir una tabla de síndromes

Paso 1: Liste todos los co-conjuntos del código, y elija para cada co-conjunto la palabra código de menor peso $\mathbf{w}$ como palabra "líder" del co-conjunto.

Paso 2: Encuentre una matriz parity check H para el código y, para cada palabra líder $\mathbf{u}$, calcule su síndrome $S(\mathbf{u}) = \mathbf{u}H^T$.

Para construir una tabla de síndromes bajo el esquema de decodificación *incompleta* por medio del vecino más cercano, si hay dos palabras o más con el mismo peso menor en el co-conjunto, se solicita una retransmisión y se indica esto en la tabla por medio de un guión $-$.

Un elemento líder único de un co-conjunto corresponde a un patrón de error que puede ser corregido, asumiendo una decodificación incompleta por medio del vecino más cercano. Un elemento líder de un co-conjunto (no necesariamente único) corresponde a un patrón de error que puede ser corregido, asumiendo una decodificación completa por medio del vecino más cercano.

1.55 Ejemplo. Puede verificarse fácilmente que

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

Elementos líder $\mathbf{u}$	Síndrome $S_H(\mathbf{u}) = \mathbf{u}H^T$
00000	00
00001	01
00010	10
11000	11

Tab. 1.1: Tabla de búsqueda de síndromes, Ejemplo 1.55.

es una matriz parity check para el código del Ejemplo 1.51. Construyamos la tabla de búsqueda de síndromes para C.

De cada co-conjunto que obtuvimos elijamos un elemento líder, es decir, la palabra del co-conjunto de menor peso (si hubiese un empate la elección es indistinta), digamos 00000, 00001, 00010 y 00011, y calculemos el síndrome de cada elemento líder haciendo $S_H(\mathbf{u}) = \mathbf{u}H^T$, siendo $\mathbf{u}$ cada elemento líder.

Finalmente, se obtiene la tabla de búsqueda de síndromes deseada, misma que se aprecia en la Tabla 1.1.

Una manera más rápida de construir una tabla de síndromes, dada una matriz parity check H y distancia d para el código C, es generar todos los posibles patrones de error $\mathbf{e}$ con

$$wt(\mathbf{e}) \leq \left\lfloor \frac{d-1}{2} \right\rfloor$$

como elementos líder y calcular el síndrome $S(\mathbf{e})$ para cada uno de ellos.

Finalmente, detallaremos cómo realizar la decodificación por síndromes.

Pasos en la decodificación por síndromes

Paso 1: Para la palabra recibida $\mathbf{w}$, calcule el síndrome $S(\mathbf{w})$.

Paso 2: Encuentre el elemento líder $\mathbf{u}$ de cada co-conjunto próximo al síndrome $S(\mathbf{w}) = S(\mathbf{u})$ en la tabla de síndromes.

Paso 3: Decodifique $\mathbf{w}$ como $\mathbf{w} - \mathbf{u}$.

1.56 Ejemplo. Sea C el código lineal binario con matriz parity check

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

Elementos líder $\mathbf{u}$	Síndrome $S_H(\mathbf{u}) = \mathbf{u}H^T$
000000	000
000001	001
000010	010
000100	100
001000	011
010000	101
100000	110
001100	111

Tab. 1.2: Tabla de búsqueda de síndromes, Ejemplo 1.56.

Haciendo un análisis análogo al que se explica en el Ejemplo 1.55, y tomando en cuenta que H se encuentra ya en su forma estándar, tenemos que la matriz generadora del código C es

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

esto es $\{100110, 010101, 001011\}$ es una base para C. Puesto que se trata de un código binario, tenemos un total de $2^3 = 8$ elementos en el código[5]:

$$C = \{000000, 100110, 010101, 001011, 110011, 101101, 011110, 111000\}.$$

Por inspección, se observa que $d(C) = 3$, y

$$wt(\mathbf{e}) \leq \left\lfloor \frac{3-1}{2} \right\rfloor = 1.$$

Entonces, todos los patrones de error con peso 0 ó 1 son elementos líderes $\mathbf{u}$ de un co-conjunto, por lo que se obtienen con facilidad los primeros 7 renglones de la tabla de búsqueda de síndromes que puede apreciarse en la Tabla 1.2.

El número de co-conjuntos que pueden obtenerse para el código C es $q^{n-k} = 2^{6-3} = 8$ y, por lo tanto, existen 8 respectivos síndromes. Esto quiere decir que falta un elemento líder de un co-conjunto por identificar, y por lo tanto un síndrome, esto es, falta un renglón de la tabla de búsqueda de síndromes.

El último renglón de la tabla se encuentra fácilmente si se considera que $wt(\mathbf{e}) = 2$, puesto que ya se agotó toda posibilidad de que $wt(\mathbf{e}) \leq 1$. Puesto que $\mathbf{u}$ debe ser un elemento líder del co-conjunto al que pertenece, y considerando una decodificación

[5] Todas las combinaciones lineales de la base de C.

completa, escogemos arbitrariamente 001100 y tenemos $\mathbf{u}H^T = 111$, el último síndrome factible para completar la tabla.

Supongamos que recibimos las siguientes palabras: 110110, 011011 y 101010. Por supuesto, dichas palabras no corresponden a ninguna de las palabras que se encuentran en el código C, pero podemos decodificarlas buscando uno de los patrones de errores de la tabla de búsqueda de síndromes:

(i) $\mathbf{u} = 110110$, $S(\mathbf{u}) = \mathbf{u}H^T = 101$, el síndrome que corresponde al co-conjunto cuyo líder es 010000 y entonces 010000+110110=100110, por lo que 110110 se decodifica como $100110 \in C$;

(ii) $\mathbf{u} = 011011$, $S(\mathbf{u}) = 101$, el síndrome que corresponde al co-conjunto cuyo líder es 010000 y entonces 010000+011011=001011, por lo que 011011 se decodifica como $001011 \in C$;

(iii) $\mathbf{u} = 101010$, $S(\mathbf{u}) = 111$, el síndrome que corresponde al co-conjunto cuyo líder es 001100 y entonces 001100+101010=100110, por lo que 101010 se decodifica como $100110 \in C$;

Con esto concluimos nuestro estudio sobre nociones básicas en la teoría de códigos. Como el lector habrá podido percatarse a lo largo de este capítulo, la teoría de códigos es muy vasta y extensa. Como libros introductorios a la teoría de códigos se recomiendan los libros de Adamek (1991), Pretzel (1998), Ling y Xing (2004) y Bierbrauer (2005). Si se desea ahondar en el tema, se recomienda ampliamente el libro de van Lint (1999). Para una revisión general sobre el desarrollo de la teoría de códigos desde sus inicios y hasta nuestros días, se recomienda revisar el libro de Berlekamp (1974) y los artículos de Sudan (2001) y Calderbank (1998).

2. COTAS EN TEORÍA DE CÓDIGOS

"Dios es una esfera terrible cuyo centro está en todas partes
y cuya circunferencia en ninguna".
Blaise Pascal.

Recordemos que $C_q(n, M)$ denota un código definido sobre F_q, un subespacio de F_q^n, de longitud n y de tamaño M. Ahora, denotaremos como $C_q(n, M, d)$ a un código de longitud n, tamaño M y distancia d. Considerando un valor de n fijo el tamaño M del código es un parámetro que permite medir la eficiencia del mismo y la distancia d indica la capacidad para corregir errores del código en cuestión. Por supuesto, sería deseable que un código fuese eficiente y que a la vez permitiera corregir el mayor número de errores posible, esto es, que tanto M como d fueran tan grandes como fuese posible. Sin embargo, como veremos en breve, esto no ocurre.

Abordaremos algunas famosas cotas inferiores y superiores para el mayor valor posible de M dados valores fijos q, n y d.

2.1. El problema principal en la teoría de códigos

2.1 Definición. Sea C un código sobre F_q con parámetros (n, M, d), la *distancia mínima relativa* de C se define como $\delta(C) = (d-1)/n$.

Cabe destacar que la distancia mínima relativa de un código C se define con frecuencia en la literatura como $\delta(C) = d/n$. Sin embargo, definirla como $(d-1)/n$ en ocasiones conduce a fórmulas ingeniosas y manipulables.

2.2 Ejemplo. Denotemos como $R(C)$ a la tasa de transmisión de la información del código C, esto es, $R(C) = \frac{(\log_q M)}{n}$. Analicemos dos sencillos ejemplos:

(i) Consideremos el código q-ario $C = F_q^n$. Se observa con facilidad que $(n, M, d) =$

$(n, q^n, 1)$, esto es, $[n, k, d] = [n, n, 1]$. Y entonces:

$$R(C) = \frac{\log_q(q^n)}{n} = \frac{n}{n} = 1$$

pero

$$\delta(C) = 0.$$

Esto quiere decir que este código tiene la mayor tasa de transmisión posible, pues los n bits de cada palabra del mismo se utilizan para enviar información. Sin embargo, el código posee una distancia mínima relativa nula, lo cual revela una capacidad nula para corregir errores.

(ii) Consideremos ahora el siguiente código binario de longitud n

$$C = \{\underbrace{00\cdots0}_{n}, \underbrace{11\cdots1}_{n}\}.$$

Es claro que la dimensión de este código es $k = 1$ y que su distancia es n, por lo que C es un código con parámetros $[n, k, d] = [n, 1, n]$, i.e., $(n, M = q^k, d) = (n, 2, n)$, y entonces

$$R(C) = \frac{\log_2(2)}{n} = \frac{1}{n},$$

y

$$\delta(C) = \frac{n-1}{n}.$$

En este caso, conforme $n \to \infty$, $R(C) \to 0$ mientras que $\delta(C) \to 1$. Esto es, conforme aumenta la longitud de palabra, el código adquiere mayor capacidad para corregir errores a costa de sacrificar su eficiencia, pues su tasa de información tiende a bajar.

Ahora que hemos visto que la eficiencia y la capacidad de corregir errores en un código están estrechamente ligadas procederemos a establecer algunas definiciones.

2.3 Definición. Dados un alfabeto A de tamaño $q > 1$ y valores de n y d, denotemos $A_q(n, d)$ al mayor valor posible de M para el que existe un código (n, M, d) sobre A, i.e.,

$$A_q(n, d) = \max\{M : \text{existe un código } (n, M, d) \text{ sobre } A\}.$$

Un código (n, M, d) cuyo tamaño M es el máximo, i.e., $A_q(n, d) = M$ se llama *código óptimo*.

Es importante señalar que el valor de $A_q(n, d)$ depende sólo de q, n y d y no de A.

El problema de determinar los valores de $A_q(n, d)$ es comúnmente conocido como *el problema principal en Teoría de Códigos*.

En lugar de considerar todos los códigos nos limitaremos a estudiar códigos lineales a través de la siguiente definición.

2.4 Definición. Para q primo y valores de n y d dados, denotemos $B_q(n,d)$ al mayor tamaño posible q^k para el que existe un código $[n,k,d]$ sobre F_q. Esto es

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

Citaremos a continuación algunas propiedades sencillas de $A_q(n,d)$ y $B_q(n,d)$.

2.5 Teorema. *Sea $q \geq 2$ una potencia prima: $p^n = q$, donde p es un número primo y n un entero positivo. Entonces*

(i) $B_q(n,d) \leq A_q(n,d) \leq q^n$ *para toda* $1 \leq d \leq n$;

(ii) $B_q(n,1) = A_q(n,1) = q^n$;

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

Demostración. La primera desigualdad en (i) se sigue de las definiciones de $A_q(n,d)$ y $B_q(n,d)$, mientras que la segunda se desprende del hecho de que en todo código (n,M,d) sobre F_q, al ser un subconjunto no vacío de F_q^n, el parámetro $M \leq q^n$.

Ahora probemos (ii). Recordemos que F_q^n es el conjunto de *todos* los vectores de longitud n sobre F_q, o sea un código $(n,q^n,1)$ o, equivalentemente, $[n,n,1]$. Como por definición $q^k \leq B_q(n,d)$ y en este caso $k = n$ se tiene que $q^n \leq B_q(n,1)$, pero por (i) sabemos que $B_q(n,d) \leq q^n$, por lo que $B_q(n,1) = A_q(n,1) = q^n$.

Para (iii), sea C un código (n,M,n) sobre F_q. Puesto que las palabras en el código son de longitud n y la distancia entre dos palabras distintas es mayor o igual a n, se sigue que la distancia entre dos palabras distintas de C es de hecho n. Esto quiere decir que todas las palabras difieren entre sí en todas las coordenadas, de manera que las M palabras toman valores de a lo más q símbolos distintos en cada una de sus coordenadas (a lo más tantas palabras como elementos en el campo) y entonces

$$B_q(n,n) \leq A_q(n,n) \leq q.$$

Este código, claramente, es de dimensión $k = 1$ puesto que con una sóla palabra pueden generarse todas las demás al multiplicarla por $\lambda \in Fq$. Y entonces

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

□

En el caso de los códigos binarios, hay algunos resultados adicionales sobre $A_2(n,d)$ y $B_2(n,d)$. Antes de estudiarlos resulta conveniente abordar el concepto de código extendido. El código extendido de un código binario se obtiene añadiendo una coordenada parity check. Esta idea puede generalizarse a códigos sobre cualquier campo finito.

2.6 Definición. Para todo código C sobre F_q, el *código extendido de C*, denotado por $\bar{C}$ se define como

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

Cuando $q = 2$, la coordenada extra $-\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} c_i$ añadida a la palabra del código $(c_1, \ldots, c_n)$ se llama *coordenada parity check*.

2.7 Teorema. *Si C es un código con parámetros (n, M, d) sobre F_q, entonces $\bar{C}$ es un código cuyos parámetros son $(n+1, M, d')$ sobre F_q, con $d \leq d' \leq d+1$. Si C es lineal, también lo es $\bar{C}$ y, además,*

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

es la matriz parity check de $\bar{C}$ *siempre que H sea la matriz parity check de C.*

2.8 Ejemplo. (i) Consideremos el código binario lineal

$$C_1 = \{000, 110, 011, 101\}.$$

Sus parámetros son $[3, 2, 2]$. Puesto que

$$\begin{aligned} c_1 + c_2 + c_3 = \bar{c_4} &\in F_2 \\ 0 + 0 + 0 &= 0 \\ 1 + 1 + 0 &= 0 \\ 0 + 1 + 1 &= 0 \\ 1 + 0 + 1 &= 0 \end{aligned}$$

el código extendido de C es el código lineal

$$\bar{C_1} = \{0000, 1100, 0110, 1010\}$$

con parámetros $[4, 2, 2]$.

(ii) De manera similar, consideremos el código binario lineal

$$C_2 = \{000, 111, 011, 100\}.$$

Sus parámetros son $[3, 2, 1]$ y su código extendido es el código lineal

$$\bar{C}_2 = \{0000, 1111, 0110, 1001\}$$

con parámetros $[4, 2, 2]$.

En el ejemplo anterior se observa que la distancia mínima $d(\bar{C})$ puede alcanzar las cotas $d(C)$ y $d(C) + 1$.

2.9 Teorema. *Supongamos que d es impar.*

(i) El código binario (n, M, d) existe si y sólo si el código binario $(n + 1, M, d + 1)$ existe. Por lo tanto, si d es impar, $A_2(n + 1, d + 1) = A_2(n, d)$.

(ii) De forma similar, un código lineal binario $[n, k, d]$ existe si y sólo si existe un código binario $[n + 1, k, d + 1]$, por lo que $B_2(n + 1, d + 1) = B_2(n, d)$.

Demostración. Para (i) supongamos que C es un código lineal binario (n, M, d) donde d es impar. Entonces, por el Teorema 2.7, $\bar{C}$ es un código $(n+1, M, d')$ con $d \leq d' \leq d+1$. Sabemos que si d es impar, entonces el menor peso de las palabras $\mathbf{x} \in C$ es impar, pues $wt(\mathbf{x}) = d$. Supongamos que la palabra cuenta con un número de unos par en sus coordenadas, en cuyo caso la coordenada parity check será 0 y entonces el peso $wt(\mathbf{x}')$ será par. Ahora, supongamos que la palabra cuenta con un número de unos impar en sus coordenadas, en cuyo caso la coordenada parity check será 1 y entonces el peso $wt(\mathbf{x}')$ también será par. Esto es, $wt(\mathbf{x}')$ es par para toda $\mathbf{x}' \in \bar{C}$.

Además, $d(\mathbf{x}', \mathbf{y}')$ es par para toda $\mathbf{x}', \mathbf{y}' \in \bar{C}$, puesto que por el Corolario 1.10 y el Lema 1.11:

$$\begin{aligned} d(\mathbf{x}', \mathbf{y}') &= wt(\mathbf{x}' + \mathbf{y}') \\ &= wt(\mathbf{x}') + wt(\mathbf{y}') - 2wt(\mathbf{x}' * \mathbf{y}'). \end{aligned}$$

de manera que d' es par. Como $d \leq d' \leq d + 1$, i.e., d' se encuentra entre dos números enteros consecutivos y es par, se concluye que $d' = d + 1$.

Hasta el momento, hemos probado que si existe un código binario C (n, M, d), entonces existe un código binario $\bar{C}$ $(n + 1, M, d + 1)$.

Para probar que si existe un código binario $\bar{C}$ $(n+1, M, d+1)$, entonces existe un código binario C (n, M, d) usaremos el siguiente argumento: supongamos que existe un código

binario D con parámetros $(n+1, M, d+1)$ donde d es impar. Elijamos dos palabras $\mathbf{x}$ e $\mathbf{y}$ en D tales que $d(\mathbf{x}, \mathbf{y}) = d+1$, i.e., $\mathbf{x}$ e $\mathbf{y}$ difieren en $d+1 \geq 2$ coordenadas. Elijamos ahora una coordenada en la que $\mathbf{x}$ e $\mathbf{y}$ difieran y sea D' el código que se obtiene al borrar esta coordenada de *todas* las palabras en D. Al perderse una coordenada se tienen ahora palabras de longitud n y además la distancia ha sido decrementada en 1, por lo que ahora es simplemente d, esto es, D' es un código binario (n, M, d).

De estos resultados se sigue que $A_2(n+1, d+1) = A_2(n, d)$.

Para probar (ii) es suficiente con observar que al probar (i) si C es lineal, también lo es $\bar{C}$ y que si D es lineal, también lo es D'. □

Es importante señalar que el teorema anterior equivale a que si d es par, entonces $A_2(n, d) = A_2(n-1, d-1)$ y el caso de (ii) es análogo.

2.2. Cotas inferiores

Abordaremos el caso de dos famosas cotas en Teoría de Códigos: la cota de la esfera (para $A_q(n, d)$) y la cota de Gilbert-Varshamov (para $B_q(n, d)$).

2.2.1. La cota de la esfera

2.10 Definición. Sea A un alfabeto de tamaño q, donde $q > 1$. Para cualquier vector $\mathbf{u} \in A^n$ y cualquier entero $r \geq 0$, la *esfera de radio r y centro* $\mathbf{u}$, denotada como $S_A(\mathbf{u}, r)$, es el conjunto $\{\mathbf{v} \in A^n : d(\mathbf{u}, \mathbf{v}) \leq r\}$.

2.11 Definición. Para un entero dado $q > 1$, un entero positivo n y un entero $r \geq 0$, definimos $V_q^n(r)$ de esta forma:

$$V_q^n(r) = \begin{cases} \binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \ldots + \binom{n}{r}(q-1)^r & \text{si } 0 \leq r \leq n; \\ q^n & \text{si } n \leq r. \end{cases}$$

2.12 Lema. *Para todo entero $r \geq 0$, la* esfera de radio r en A^n *tiene* volumen $V_q^n(r)$. *Esto es, $S_A(\mathbf{u}, r)$ con $\mathbf{u} \in A^n$, contiene exactamente $V_q^n(r)$ vectores, donde A es un alfabeto de tamaño $q > 1$.*

Demostración. Fijemos un vector $\mathbf{u} \in A^n$. Determinaremos el número de vectores $\mathbf{v} \in A^n$ tales que $d(\mathbf{u}, \mathbf{v}) = m$, esto es, el número de vectores en A^n cuya distancia a $\mathbf{u}$ es exactamente m. El número de formas en que se pueden elegir m coordenadas en las

que $\mathbf{v}$ difiere de $\mathbf{u}$ está dado por $\binom{n}{m}$. Para cada coordenada, tenemos $q-1$ posibles símbolos para dicha coordenada en $\mathbf{v}$. Por lo tanto, el número total de vectores cuya distancia a $\mathbf{u}$ es m está dada por $\binom{n}{m}(q-1)^m$. Así, la primera parte del lema queda probada $(0 \leq r \leq n)$.

La segunda parte del lema claramente se cumple, pues cuando $r \geq n$, $S_A(\mathbf{u}, r) = A^n$ por lo que contiene $V_q^n(r) = q^n$ vectores. □

2.13 Ejemplo. Para cada una de las siguientes esferas en $A^n = F_2^n$, listaremos sus elementos y calcularemos su volumen:

(i) $S_A(1100, 3)$.

(ii) $S_A(110, 4)$.

(i) Por definición, $S_A(1100, 3) = \{\mathbf{v} \in A^4 = F_2^4 : d(1100, \mathbf{v}) \leq 3\}$.

Puesto que $r = 3$ y $n = 4$, el volumen de la esfera $|S_A(1100, 3)|$ está dado por $V_2^4(3)$:

$$\begin{aligned} V_2^4(3) &= \binom{4}{0} + \binom{4}{1}(2-1) + \binom{4}{2}(2-1)^2 + \binom{4}{3}(2-1)^3 \\ &= 1 + 4(1) + 6(1) + 4(1) \\ &= 15. \end{aligned}$$

Esto quiere decir que la esfera contiene 15 elementos. Comenzaremos por encontrar aquéllos que satisfacen $d(1100, \mathbf{v}) = 3$. En este caso, se cambiarán 3 de las 4 coordenadas, por ejemplo,

$$\underline{***}0.$$

Como hay un elemento fijo (el cero) hay $(q-1)$ posibilidades para cada coordenada $\underline{*}$, o sea, $(q-1)^3 = 1$ y esto puede ocurrir en $\binom{4}{3} = 4$ formas distintas. De aquí se tiene el último término de $V_2^4(3)$. Concretamente, los elementos que satisfacen $d(1100, \mathbf{v}) = 3$ son $\underline{001}0, \underline{0}1\underline{11}, 1\underline{011}$ y $\underline{00}0\underline{1}$.

Haciendo un análisis análogo, los elementos que satisfacen $d(1100, \mathbf{v}) = 2$ son 6:

$$\underline{00}00, \underline{0}1\underline{1}0, \underline{0}1\underline{0}\underline{1}, 1\underline{01}0, 1\underline{0}0\underline{1}\text{y } 1\underline{11}1;$$

los que satisfacen $d(1100, \mathbf{v}) = 1$ son 4:

$$\underline{0}100, 1\underline{0}00, 11\underline{1}0\text{y } 110\underline{1};$$

y el que satisface $d(1100, \mathbf{v}) = 0$ es, naturalmente, 1100.

(ii) $S_A(110, 4) = \{\mathbf{v} \in A^3 = F_2^3 : d(110, \mathbf{v}) \leq 4\}$, y como $r = 4 = n$ el volumen de esta esfera está dado por:

$$\begin{aligned} V_2^3(4) &= \binom{3}{0} + \binom{3}{1}(2-1) + \binom{3}{2}(2-1)^2 + \binom{3}{3}(2-1)^3 \\ &= 8 \\ &= q^n. \end{aligned}$$

Utilizando el mismo razonamiento que en (i), el elemento que satisface $d(110, \mathbf{v}) = 3$ es $\underline{001}$; los que satisfacen $d(110, \mathbf{v}) = 2$ son 3: $\underline{00}0, \underline{0}1\underline{1}$ y $1\underline{01}$; los que satisfacen $d(110, \mathbf{v}) = 1$ son 3: $\underline{0}10, 1\underline{0}0$ y $11\underline{1}$; y finalmente, el que satisface $d(110, \mathbf{v}) = 0$ es el centro de la esfera 110.

2.14 Teorema (Cota de la esfera)**.** *Para un entero $q > 1$ y enteros, n, d tales que $1 \leq d \leq n$, tenemos que*

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

Demostración. Sea $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\}$ un código óptimo (n, M, d) sobre A con $|A| = q$, o sea $M = A_q(n, d)$. Puesto que C tiene el máximo tamaño, no puede haber palabra en A^n cuya distancia a alguna palabra en C sea mayor o igual a d:

$$\begin{aligned} d &> d(\mathbf{x}, \mathbf{c}_i), \qquad \mathbf{x} \in A^n, \mathbf{c}_i \in C, \\ d - 1 &\geq d(\mathbf{x}, \mathbf{c}_i). \end{aligned}$$

Esto significa que para todo vector $\mathbf{x} \in A^n$ existe al menos una palabra $\mathbf{c}_i \in C$ tal que $d(\mathbf{x}, \mathbf{c}_i)$ es a lo más $d-1$, esto es $\mathbf{x} \in S_A(\mathbf{c}_i, d-1)$. Por lo tanto, toda palabra en A^n se encuentra en al menos una de las esferas $S_A(\mathbf{c}_i, d-1)$:

$$A^n \subseteq \bigcup_{i=1}^{M} S_A(\mathbf{c}_i, d-1).$$

Como $|A^n| = q^n$ y $|S_A(\mathbf{c}_i, d-1)| = V_q^n(d-1)$ para cualquier i tenemos que

$$q^n \leq M \cdot V_q^n(d-1),$$

lo cual implica que

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

□

En algunos casos especiales es posible determinar con facilidad los valores de $A_q(n,d)$, tal como se muestra en el siguiente ejemplo.

2.15 Ejemplo. Probaremos que $A_2(5,4) = 2$.

La cota de la esfera muestra que $A_2(5,4) \geq 2$ y por el Teorema 2.9 $A_2(5,4) = A_2(4,3)$, de manera que buscaremos probar que $A_2(4,3) \leq 2$. Sea C un código binario $(4, M, 3)$ y sea (x_1, x_2, x_3, x_4) una palabra en C. Dado que $d(C) = 3$, las demás palabras en C deben ser de alguna de las siguientes formas:

$$(x_1, \overline{x_2}, \overline{x_3}, \overline{x_4}), (\overline{x_1}, x_2, \overline{x_3}, \overline{x_4}), (\overline{x_1}, \overline{x_2}, x_3, \overline{x_4}),$$
$$(\overline{x_1}, \overline{x_2}, \overline{x_3}, x_4), (\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}),$$

donde $\overline{x_i}$ se define como

$$\overline{x_i} = \begin{cases} 1 & \text{si } x_i = 0; \\ 0 & \text{si } x_i = 1. \end{cases}$$

Pero, al analizar cada una de las formas de esta cinco palabras propuestas se observa que ninguna pareja tiene distancia 3 (o mayor) entre sí, de manera que sólo una de ellas puede ser incluida en el código C. Por tanto, $M \leq 2$, lo cual implica que $A_2(4,3) \leq 2$.

Como $2 \leq A_2(4,3) \leq 2$ se concluye que $A_2(5,4) = A_2(4,3) = 2$.

2.2.2. La cota de Gilbert-Varshamov

La cota de Gilbert-Varshamov es una cota inferior para $B_q(n,d)$, i.e., para códigos lineales conocida desde los años 50. Existe una cota asintótica de Gilbert-Varshamov, misma que durante muchos años fue considerada la mejor cota inferior por sus alcances para familias infinitas de códigos lineales, por lo que llegó a convertirse en una clase de estándar para juzgar la "bondad"de una secuencia infinita de códigos lineales. Entre 1977 y 1982, V.D. Goppa construyó códigos con geometrías algebraicas utilizando curvas algebraicas sobre campos finitos con varios puntos racionales (el artículo de Berlekamp (1973) hace una revisión sobre los artículos de Goppa). Se logró un gran avance en la teoría de códigos poco después de dicho descubrimiento, cuando se mostró que existían secuencias de códigos con geometrías algebraicas que se comportaban mejor que las cotas asintóticas de Gilbert-Varshamov para ciertas q lo suficientemente grandes.

2.16 Teorema (Cota de Gilbert-Varshamov)**.** *Sean n, k y d enteros que satisfacen $2 \leq d \leq n$ y $1 \leq k \leq n$. Si*

$$\sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k}, \tag{2.1}$$

entonces existe un código con parámetros $[n,k]$ sobre F_q con distancia mínima mayor o igual a d.

Demostración. Probaremos que, si (2.1) se satisface, entonces existe una matriz H de $(n-k) \times n$ sobre F_q tal que todas las $d-1$ columnas son linealmente independientes.

Construiremos H como sigue. Sea $\mathbf{c}_j$ la j-ésima columna de H.

Sea $\mathbf{c}_1$ cualquier vector distinto de cero en F_q^{n-k}. Sea $\mathbf{c}_2$ cualquier vector que no se encuentre en la expansión de $\mathbf{c}_1$, esto es, que no pueda ser generado por medio de alguna combinación lineal de $\mathbf{c}_1$. Para cualquier $2 \leq j \leq n$, sea $\mathbf{c}_j$ cualquier vector que no se encuentra en la expansión lineal de $d-2$ (o menos) vectores $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{j-1}$.

El número de vectores en la expansión lineal de $d-2$ (o menos) de las palabras $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$ $(2 \leq j \leq n)$ está dada por

$$\sum_{i=0}^{d-2} \binom{j-1}{i} (q-1)^i \leq \sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k}.$$

Por lo tanto, el vector $\mathbf{c}_j$ $(2 \leq j \leq n)$ siempre puede encontrarse.

La matriz H que se construye de esta forma es una matriz de $(n-k) \times n$ y cualesquiera $d-1$ columnas de H son linealmente independientes. El espacio nulo de H es un código lineal sobre F_q de longitud n, distancia al menos d, y dimensión al menos k. Para obtener un código lineal de las características deseadas, simplemente consideremos un subespacio k dimensional. □

2.3. Cotas superiores

A continuación presentaremos algunas de las cotas superiores más famosas en la teoría de códigos.

2.3.1. La cota de Hamming

La siguiente cota fue hallada por Richard W. Hamming en 1950.

2.17 Teorema (Cota de Hamming)**.** *Para todo entero $q > 1$ y enteros n, d tales que $1 \leq d \leq n$, se satisface que*

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

Demostración. Sea $C = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M)$ un código óptimo de parámetros (n, M, d) sobre A (con $|A| = q$), de manera que $M = A_q(n, d)$. Sea $e = \left\lfloor \frac{(d-1)}{2} \right\rfloor$; entonces, las esferas $S_A(\mathbf{c_i}, e)$ son disjuntas. Por esta razón, tenemos que

$$\bigcup_{i=1}^{M} S_A(\mathbf{c_i}, e) \subseteq A^n,$$

donde la unión de esferas es disjunta. Puesto que $|A^n = q^n|$ y $|S_A(\mathbf{c_i}, e)| = V_q^n(e)$ para cualquier i, se cumple que

$$M \cdot V_q^n(e) \leq q^n,$$

lo cual implica que

$$A_q(n, d) = M \leq \frac{q^n}{V_q^n(e)} = \frac{q^n}{V_q^n\left(\frac{\lfloor (d-1) \rfloor}{2}\right)}$$

completando así la prueba. □

2.18 Definición. Un código q-ario que alcanza la cota de Hamming, esto es, aquel que tiene

$$\frac{q^n}{\sum_{i=o}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i}(q-1)^i}$$

palabras, se llama *código perfecto.*

Los siguientes son casos simples de códigos perfectos, por lo que se conocen como *códigos perfectos triviales*:

(i) el código lineal $C = F_q^n$, pues en este caso $d = 1$ y entonces

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

y $|C| = q^n$;

(ii) todo C con $|C| = 1$, pues en este caso $d = \infty$ y entonces

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

y $|C| = 1$ $(k = 1)$;

(iii) códigos binarios repetidos de longitud impar consistentes en 2 palabras código con distancia n la una de la otra ($d = n$). Puesto que n es impar, $n - 1$ es par, entonces

$$\left\lfloor \frac{n-1}{2} \right\rfloor = \frac{n-1}{2}$$

.

El valor de k que maximiza $\binom{n}{k}$, para una n fija, es

$$k = \left\lfloor \frac{n}{2} \right\rfloor = \left\lfloor \frac{n-1}{2} + \frac{1}{2} \right\rfloor = \frac{n-1}{2},$$

lo que equivale a sumar sólo la mitad de elementos que aparecen en el renglón n del triángulo de Pascal. Como sabemos, la suma de cada renglón del triángulo de Pascal es una potencia de 2, por lo que se tiene:

$$\sum_{i=0}^{\frac{n-1}{2}} \binom{n}{i} = \frac{1}{2} \sum_{i=0}^{n} \binom{n}{i} = \frac{1}{2} 2^n = 2^{n-1}.$$

Entonces, la cota de Hamming está dada por:

$$\frac{2^n}{\sum_{i=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} \binom{n}{i} (2-1)^i} = \frac{2^n}{2^{n-1}} = 2,$$

y $|C| = 2$, $(k = 1)$.

Dos ejemplos no triviales de códigos perfectos son los códigos de Hamming y los códigos de Golay. Las características y propiedades de dichos códigos se discuten de forma clara en el libro de Ling y Xing (2004), en el que se abordan los casos binarios y ternarios para posteriormente explorar la generalización al caso q-ario.

De hecho un importante resultado obtenido por Tietavainen (1973), señala que un código perfecto no trivial definido sobre F_q, en el que $q \geq 2$ es una potencia prima, debe tener los mismos parámetros que un código de Golay o uno de Hamming.

Casi una década después de que Hamming propusiera esta cota, Johnson (1962) planteó una extensión de la de Hamming bajo argumentos combinatorios. La cota superior de Hamming se estima para códigos correctores de e errores considerando todos aquellos puntos tales que $d(\mathbf{c}, C) \leq e$, $\mathbf{c} \in C$, mientras que la de Johnson abarca también los casos en que $d(\mathbf{c}, C) > e$. A pesar de que en los casos $5 \geq e$ la cota de Johnson no permite encontrar una cota superior para $A_2(n, 2e + 1)$, Haas (2008) encontró una técnica para ello, cerrando así una brecha más en la teoría de códigos.

Para ahondar en el tema de los códigos perfectos, se recomienda el artículo de van Lint (1975).

2.3.2. La cota de Singleton

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

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

Particularmente, cuando q es una potencia prima, los parámetros $[n, k, d]$ de cualquier código lineal sobre F_q satisfacen

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

Demostración. Consideremos un código con parámetros (n, M, d) sobre el alfabeto A de tamaño q, donde $M = A_q(n, d)$. Denotemos $\mathbf{x}^{(i)}$, a cada palabra en C. Eliminemos las últimas $d - 1$ coordenadas de cada palabra en C:

$$\mathbf{x}^{(i)} = \overbrace{\underbrace{x_1 \cdots x_{n-d+1}}_{n-(d-1)=n-d+1} \underbrace{x_{n-d+2} \cdots x_n}_{d-1}}^{n}.$$

Como la distancia de C es d, al eliminar $d - 1$ coordenadas en todas sus palabras se tiene un código de longitud $n - d + 1$, y todas las palabras código son distintas entre sí ya que la distancia es, por lo menos, 1.[1]

El número máximo de palabras de longitud $n-d+1$ es q^{n-d+1}, esto es, $A_q(n, d) = M \leq q^{n-d+1}$.

Además, dada cualquier matriz parity check H para C, el rango de H es por definición $n - k$ y, por tanto, cualesquiera $n - k + 1$ columnas de H son linealmente dependientes.

Por el Teorema 1.42 sabemos que C tiene distancia $\leq m$ si y sólo si H tiene m columnas linealmente dependientes. Por lo que $d \leq n - k + 1$, esto es, $d + k \leq n + 1$, probando así la última parte del teorema. □

Una forma alternativa de establecer la cota de Singleton es la siguiente. Para cualquier

[1] ¡El número de palabras en el código sigue siendo el mismo!.

código q-ario C se cumple que[2]

$$\begin{aligned} R(C) + \delta(C) &\leq 1 \\ \frac{\log_q |C|}{n} + \frac{d-1}{n} &\leq 1 \\ \frac{\log_q q^k}{n} + \frac{d-1}{n} &\leq 1 \\ \frac{k+d-1}{n} &\leq 1 \\ d + k &\leq n+1. \end{aligned}$$

Cuando un código alcanza la cota de Singleton, se tiene un código DMS.

2.20 Definición. Un código lineal con parámetros $[n, M, d]$ tal que $k + d = n + 1$ se llama *código de distancia mínima separable* (DMS).

En su artículo, *Maximum distance q-nary codes*, Singleton (1964) incluyó ejemplos de estos códigos y explicó métodos para construirlos y probar que existen códigos DMS para valores de q, k y r.

2.3.3. La cota de Plotkin

Esta es una cota superior propuesta por Plotkin (1960) y se distingue por ser una de las cotas más restrictivas encontradas a la fecha. Para comprender mejor esta cota conviene recordar la famosa desigualdad de Cauchy-Schwarz en el siguiente lema. Para revisar la demostración detallada del mismo, se recomienda al lector el libro Análisis Matemático Vol. 1 de Haaser y otros (1972).

2.21 Lema (Desigualdad de Cauchy-Schwarz)**.** *Sean* $\{a_1, \ldots, a_n\}$ *y* $\{b_1, \ldots, b_n\}$ *dos conjuntos de números reales. Entonces*

$$\left(\sum_{r=1}^{m} a_r b_r\right)^2 = \left(\sum_{r=1}^{m} a_r^2\right)\left(\sum_{s=1}^{m} b_s^2\right) - \sum_{r=1}^{m}\sum_{s=1}^{m} \left(a_r b_s - a_s b_r\right)^2 /2$$

Por lo que

$$\left(\sum_{r=1}^{m} a_r b_r\right)^2 \leq \left(\sum_{r=1}^{m} a_r^2\right)\left(\sum_{r=1}^{m} b_r^2\right). \tag{2.2}$$

[2] Véanse las primeras páginas del presente capítulo.

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

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

Demostración. Sea $C - (n, M, d)$ sobre un alfabeto A de tamaño q, y sea

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

la distancia entre cada pareja de palabras en el código. Puesto que d es minimal, $d \leq d(\mathbf{c}, \mathbf{c}')$ para $\mathbf{c}$ y $\mathbf{c}' \in C$ tales que $\mathbf{c} \neq \mathbf{c}'$ y se sigue que

$$\begin{aligned} \sum_{\mathbf{c}' \in C} d &\leq \sum_{\mathbf{c}' \in C} d(\mathbf{c}, \mathbf{c}') \\ (M-1)d &\leq \sum_{\mathbf{c}' \in C} d(\mathbf{c}, \mathbf{c}') \\ \sum_{\mathbf{c} \in C} (M-1)d &\leq \sum_{\mathbf{c} \in C} \sum_{\mathbf{c}' \in C} d(\mathbf{c}, \mathbf{c}') \\ M(M-1)d &\leq \sum_{\mathbf{c} \in C} \sum_{\mathbf{c}' \in C} d(\mathbf{c}, \mathbf{c}'). \end{aligned}$$

Esto es

$$M(M-1)d \leq T. \tag{2.3}$$

Ahora, sea A un arreglo de $M \times n$, es decir, una lista cuyos renglones son las M palabras en C, mismas que son de longitud n. Para $1 \leq i \leq n$ y $a \in A$, sea $n_{i,a}$ el número de entradas en la i-ésima columna de A que son iguales a a.

Se tiene que

$$\sum_{a \in A} n_{i,a} = M$$

para toda $1 \leq i \leq n$, ya que hay M renglones en la lista y si sumamos el número de veces que aparece cada símbolo en la i-ésima columna en precisamente M.

Sean $\mathbf{c} = (c_1, \ldots, c_n)$ y $\mathbf{c}' = (c'_1, \ldots, c'_n)$, de manera que ahora podemos escribir T considerando la distancia símbolo por símbolo

$$T = \sum_{i=1}^{n} \left(\sum_{\mathbf{c} \in C} \sum_{\mathbf{c}'} d(c_i, c'_i) \right).$$

Además,

$$\sum_{a\in A} n_{i,a}(M - n_{i,a})$$

es la contribución de la i -ésima columna a la suma de las distancias entre todos los pares ordenados de palabras distintas entre sí, por lo que entonces

$$\begin{aligned}
T &= \sum_{i=1}^{n}\sum_{a\in A} n_{i,a}(M - n_{i,a}) \\
&= \sum_{i=1}^{n}\sum_{a\in A} \left(Mn_{i,a} - n_{i,a}^2\right) \\
&= \sum_{i=1}^{n}\left(\sum_{a\in A} Mn_{i,a} - \sum_{a\in A} n_{i,a}^2\right) \\
&= \sum_{i=1}^{n}\left(M\sum_{a\in A} n_{i,a} - \sum_{a\in A} n_{i,a}^2\right) \\
&= \sum_{i=1}^{n} M^2 - \sum_{i=1}^{n}\sum_{a\in A} n_{i,a}^2 \\
&= nM^2 - \sum_{i=1}^{n}\sum_{a\in A} n_{i,a}^2.
\end{aligned}$$

Aplicando la desigualdad de Cauchy-Schwarz haciendo $m = q$ y $a_1 = \ldots = a_q = 1$, se sigue que

$$\left(\sum_{a\in A} n_{i,a}\right)^2 \leq \left(\sum_{i=1}^{q} 1^2\right)\left(\sum_{a\in A} n_{i,a}^2\right) = q\sum_{a\in A} n_{i,a}^2.$$

De donde se tiene que

$$\begin{aligned}
q\sum_{a\in A} n_{i,a}^2 &\geq \left(\sum_{a\in A} n_{i,a}\right)^2 \\
\sum_{a\in A} n_{i,a}^2 &\geq q^{-1}\left(\sum_{a\in A} n_{i,a}\right)^2 \\
-\sum_{a\in A} n_{i,a}^2 &\leq -q^{-1}\left(\sum_{a\in A} n_{i,a}\right)^2 \\
-\sum_{i=1}^{n}\sum_{a\in A} n_{i,a}^2 &\leq -\sum_{i=1}^{n} q^{-1}\left(\sum_{a\in A} n_{i,a}\right)^2.
\end{aligned}$$

Y entonces

$$nM^2 - \sum_{i=1}^{n} \sum_{a \in A} n_{i,a}^2 \leq nM^2 - \sum_{i=1}^{n} q^{-1} \left(\sum_{a \in A} n_{i,a} \right)^2 .$$

Esto es

$$T \leq nrM^2. \tag{2.4}$$

La cota de Plotkin se sigue de (2.3) y (2.4) por transitividad:

$$\begin{aligned} (M-1)d &\leq nrM \\ dM - d &\leq nrM \\ dM - nrM &\leq d \\ M(d-nr) &\leq d \\ M &\leq \frac{d}{d-nr}. \end{aligned}$$

Puesto que $d/(d-nr)$ puede ser real y M es estrictamente un entero positivo se discretiza la cota

$$M \leq \left\lfloor \frac{d}{d-nr} \right\rfloor . \tag{2.5}$$

El último paso es válido ya que sabemos que $n \leq \lfloor x \rfloor$ si, y sólo si, $n \leq x$. □

En el caso particular de códigos sobre F_2 se tiene una versión más refinada y útil de la cota de Plotkin. Antes de establecer la cota de Plotkin para códigos binarios es conveniente revisar el siguiente teorema que será de gran utilidad para este fin.

2.23 Teorema. $A_q(n,d) \leq qA_q(n-1,d)$.

Demostración. Sea C el código óptimo de parámetros (n, M, d) sobre F_q. Sea $C = \bigcup_{i=0}^{q-1} C^{(i)}$, donde $C^{(i)}$ es el conjunto de palabras cuya última coordenada es i. Entonces, para toda $0 \leq i \leq q-1$,

$$\left| \bigcup_{i=0}^{q-1} C^{(i)} \right| = q \left| C^{(i)} \right| .$$

Además cada $C^{(i)} \leq A_q(n-1,d)$. De donde se sigue que $A_q(n,d) \leq qA_q(n-1,d)$. □

2.24 Teorema (Cota de Plotkin para códigos binarios)**.** (i) *Cuando d es par,*

$$A_2(n,d) \leq \begin{cases} 2\lfloor d/(2d-n) \rfloor & \text{para } n < 2d; \\ 4d & \text{para } n = 2d. \end{cases}$$

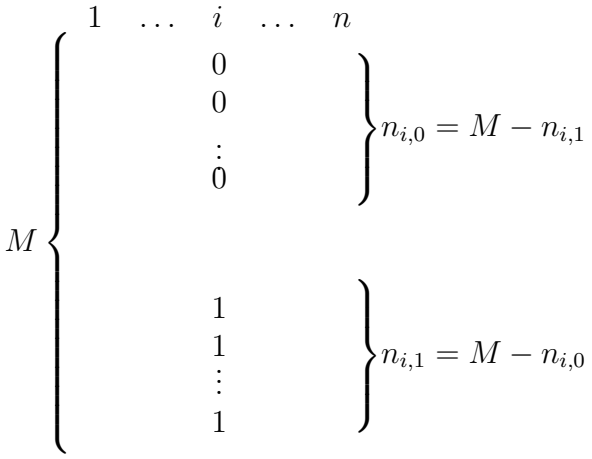


Fig. 2.1: $n_{i,a}$ en un código binario.

(ii) *Cuando d es impar*

$$A_2(n,d) \leq \begin{cases} 2\lfloor (d+1)/(2d+1-n) \rfloor & \textit{para } n \leq 2d+1, \\ 4d+4 & n{=}2d{+}1. \end{cases}$$

Demostración. (i) Primero, analicemos la cota en el caso binario para M par e impar. Cuando $q = 2$ se tiene que

$$\sum_{i=1}^{n} \sum_{a \in A} n_{i,a}(M - n_{i,a}) = \sum_{i=1}^{n} 2n_{i,a}(M - n_{i,a}) \tag{2.6}$$

lo cual se muestra en la Figura 2.1.

Cuando M es par, el máximo valor que toma $n_{i,a}(M - n_{i,a})$ está dado por

$$max(n_{i,a}(M - n_{i,a})) = \{n_{i,a} = M - n_{i,a} : M \text{ es par}\}$$

lo cual ocurre en este caso cuando $n_{i,a} = M/2$ o, en otras palabras, cuando hay tantos unos como ceros en la i-ésima columna.

Y entonces, T queda expresada como

$$\sum_{i=1}^{n} M \left(M - \frac{M}{2} \right) = \frac{nM^2}{2}.$$

Esto es, $M(M-1)d \leq T \leq (nM^2)/2$. Y se tiene por transitividad, para el caso de M par, $\frac{M}{2} \leq \frac{d}{2d-n}$ lo cual ocurre si y sólo si, $M/2 \leq \lfloor d/(2d-n) \rfloor$, es decir, si y sólo si

$$M \leq 2\lfloor d/(2d-n) \rfloor, \quad 2d - n > 0. \tag{2.7}$$

Note que $2d - n < 0$ no es factible en virtud de que M se considera siempre un entero positivo.

Por otra parte, cuando M es impar, el máximo valor de $n_{i,a}(M - n_{i,a})$ está dado por

$$max(n_{i,a}(M - n_{i,a})) = \{n_{i,a} = M - n_{i,a} \pm 1 : M \text{ es impar}\}$$

es decir, cuando $n_{i,a} = (M \pm 1)/2$, en cuyo caso se tiene para T que

$$\begin{aligned}\sum_{i=1}^{n}(M \pm 1)\left(M - \frac{M \pm 1}{2}\right) =& \sum_{i=1}^{n}\frac{(M \pm 1)(M \mp 1)}{2} \\ =& \frac{n(M^2 - 1)}{2}.\end{aligned}$$

Por lo tanto, $M(M-1)d \leq n(M^2-1)/2 = (nM^2-n)/2 \leq nM^2$, y desarrollando se tiene nuevamente (2.7), ahora para M impar.

Esto significa que sin importar si el número máximo de palabras en un código binario es par o impar, se alcanza la primera cota propuesta, lo cual prueba el caso d par y $n < 2d$.

Para el caso en el que d es par y $n = 2d$ consideremos que, por el Teorema 2.23,

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

Calculemos la cota para $A_2(n - 1, d)$. En dicho caso T se expresa como

$$\sum_{i=1}^{n-1}\frac{M^2}{2} = \frac{(n-1)M^2}{2}.$$

Entonces $M(M - 1)d \leq T \leq ((n - 1)M^2)/2$, de donde se sigue que $M/2 \leq d/(2d - n + 1)$, lo cual ocurre si y sólo si $M \leq \lfloor d/(2d - n + 1)\rfloor = 2d$, siendo $n = 2d$.

Ahora bien, como $A_2(n, d) \leq 2d$ y $A_2(n, d) \leq 2A_2(n - 1, d) \leq 4d$, $A_2(n, d) \leq 4d$.

(ii) Para probar la segunda parte del teorema recordemos que $A_2(n, d) = A_2(n+1, d+1)$ para d impar (resultado (ii) del Teorema 2.9), por lo que podemos calcular la cota para $A_2(n + 1, d + 1)$, misma que también es válida para $A_2(n, d)$. Por tal motivo calculamos T de la siguiente manera:

$$T = \sum_{i=1}^{n-1}\frac{M^2}{2} = \frac{(n+1)M^2}{2},$$

y entonces $2M(M - 1)(d + 1) \leq (n + 1)M^2$, y simplificando

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

esto es,

$$M \leq 2 \left\lfloor \frac{d+1}{2d+1-n} \right\rfloor, \qquad 2d+1 > n.$$

El último caso, $n = 2d+1$ y d impar, se desprende de la desigualdad $A_2(n,d) = A_2(n+1, d+1) \leq 2A_2(n, d+1)$, misma que se obtiene de combinar el Teorema 2.9 (i) y el Teorema 2.23. La cota para $A_2(n, d+1)$ en este caso es $nM^2/2$ y se tiene que $M(M-1)(d+1) \leq (nM^2)/2$. Simplificando se obtiene

$$\frac{M}{2} \leq \frac{d+1}{2d+2-n}$$

y sustituyendo n y discretizando la cota $M/2 \leq d+1$.

Como

$$A_2(n, d+1) \leq 2d+2,$$

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

Y por transitividad $A_2(n,d) \leq 4d+4$, completando la demostración del teorema.

□

El trabajo de Plotkin fue retomado hace escasamente 5 años por Quistorff (2003), quien mejoró la cota de Plotkin para códigos construidos sobre la métrica de Lee. [3]

2.3.4. La cota de Griesmer

La última cota que estudiaremos en este capítulo es la cota (superior) de Griesmer, misma que es válida únicamente para códigos lineales. Para una mejor comprensión de esta última cota, estableceremos algunas definiciones preliminares. Sea C un código lineal sobre F_q con parámetros $[n,k]$ y supongamos que $\mathbf{c}$ es una palabra código en C con peso $wt(\mathbf{c}) = w$.

2.25 Definición. El *soporte* de $\mathbf{c}$, denotado por $Sop(\mathbf{c})$, es el conjunto de coordenadas de $\mathbf{c}$ distintas a cero.

2.26 Definición. El *código residual* de C con respecto de $\mathbf{c}$, denotado como $Res(C, \mathbf{c})$, es el código de longitud $n-w$ obtenido de C al eliminar todas las coordenadas de $Sop(C)$.

[3] Recuérdese que, al igual que en capítulo anterior, en este capítulo los códigos se construyen bajo la métrica de Hamming (espacio Hamminiano).

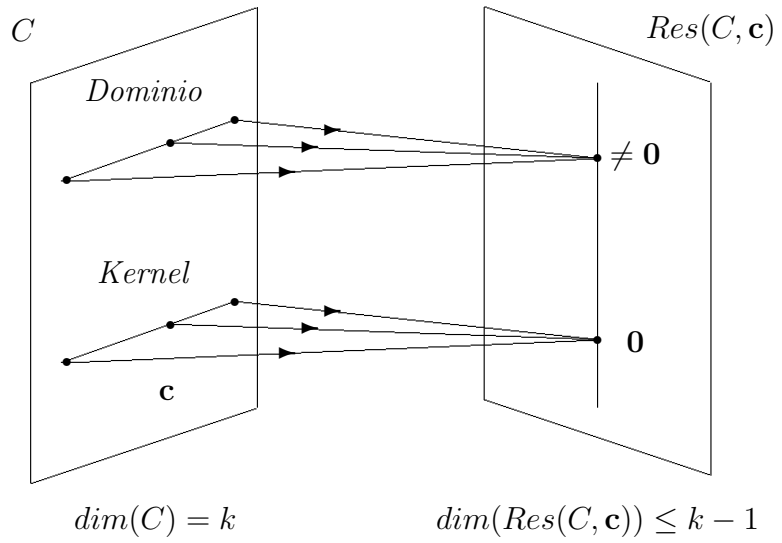


Fig. 2.2: Kernel (o espacio nulo) de una transformación lineal

Note que $w = |Sop(\mathbf{c})|$.

2.27 Ejemplo. Sea $C = \{101, 001, 110\}$. Consideremos $\mathbf{c} = 101$, entonces $Sop(\mathbf{c}) = \{c_1, c_3\}$, las coordenadas 1 y 3 de $\mathbf{c}$. $wt = w = |Sop(\mathbf{c})| = 2$, por lo que la longitud del código residual obtenido a partir de C es $n-w = 3-2 = 1$ y entonces $Res(C, \mathbf{c}) = \{0, 1\}$.

2.28 Definición. Una función $f : A \to B$ se llama *suprayectiva* si para toda $b \in B$ existe $a \in A$ tal que $f(a) = b$.

2.29 Definición. Una *transformación lineal* T de un espacio vectorial V en un espacio vectorial W es una regla que asigna a cada vector $\mathbf{x}$ en V un vector único en $T(\mathbf{x})$ en W, tal que

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, para toda $\mathbf{u}, \mathbf{v} \in V$; y

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$, para toda $\mathbf{u} \in V$ y todo escalar c.

2.30 Definición. El *kernel* (o espacio nulo) de una transformación lineal es el conjunto de todas las $\mathbf{u} \in V$ tales que $T(\mathbf{u}) = \mathbf{0}$.

Por último, estableceremos un resultado que nos será de mucha utilidad conocido como el *principio del palomar.*

2.31 Proposición (Principio del palomar generalizado)**.** *Si m palomas ocupan n palomares y $m > kn$, donde k es un entero positivo, entonces* al menos *un palomar tiene $k + 1$ palomas.*

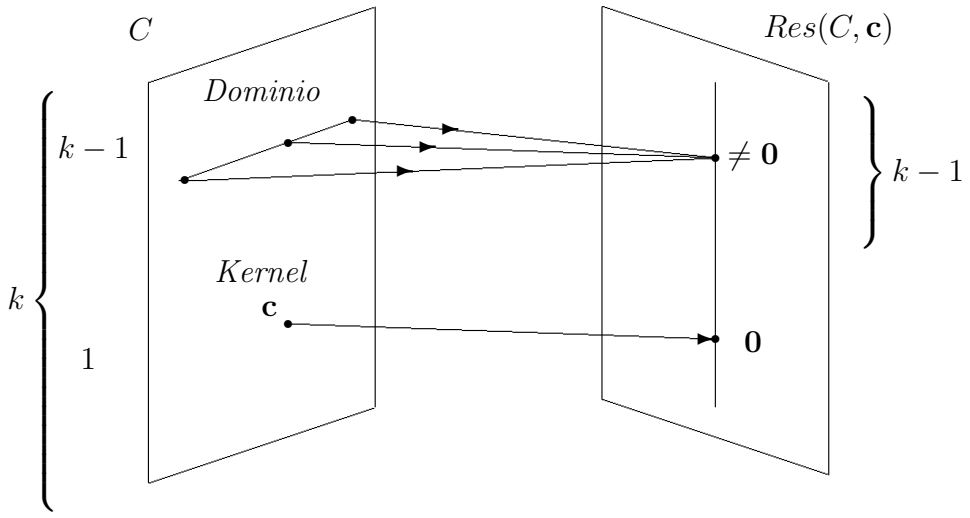


Fig. 2.3: Kernel de una transformación lineal: $dim(Res(C, \mathbf{c})) = k - 1$.

2.32 Lema. *Si C es un código $[n, k, d]$ sobre F_q y $\mathbf{c} \in C$ es una palabra código de peso d, entonces $Res(C, \mathbf{c})$ es un código $[n - d, k - 1, d']$, donde $d' \geq \lceil d/q \rceil$.*

Demostración. Sin perder generalidad, podemos reemplazar C por un código equivalente de manera que $\mathbf{c} = (1, 1, \ldots, 1, 0, 0, \ldots, 0)$ donde las primeras d coordenadas son 1; y las demás, ceros.

$Res(C, \mathbf{c})$ tiene dimensión a lo más de $k - 1$. Observemos que $Res(C, \mathbf{c})$ es un código lineal. Para toda $\mathbf{x} \in F_q^n$, denotamos por $\mathbf{x}'$ el vector obtenido de $\mathbf{x}$ al borrar las primeras d coordenadas, esto es, al eliminar las coordenadas de $Sop(\mathbf{c})$. Ahora, es fácil ver que el mapeo $C \to Res(C, \mathbf{c})$ dado por $\mathbf{x} \mapsto \mathbf{x}'$ es una transformación lineal de espacios vectoriales cuyo kernel contiene a $\mathbf{c}$ y es entonces un subespacio de C de dimensión al menos 1, por lo que $Res(C, \mathbf{c}')$ tiene dimensión de a lo más $k - 1$ (véase la Figura 2.3).

Ahora, mostraremos que $Res(C, \mathbf{c})$ tiene dimensión $k-1$. Supongamos que la dimensión es estrictamente menor a $k - 1$. Entonces hay una palabra $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ en C que no es múltiplo de $\mathbf{c}$ y que tiene la propiedad de que $v_{d+1} = \ldots = v_n = 0$. Entonces $\mathbf{v} - v_1\mathbf{c}$ es una palabra distinta de $\mathbf{0}$ y que pertenece a C:

$$\begin{aligned} \mathbf{v} - v_1\mathbf{c} &= (v_1, v_2, v_3, \ldots, v_d, v_{d+1}, \ldots, v_n) - v_1(1, 1, 1, \ldots, 1, 0, \ldots, 0) \\ &= (0, \underbrace{v_2 - v_1, v_3 - v_1, \ldots, v_d - v_1}_{wt(\mathbf{v} - v_1\mathbf{c}) \leq d-1 < d}, 0, \ldots, 0). \end{aligned}$$

contradiciendo la definición de d, por lo que $Res(C, \mathbf{c})$ tiene dimensión $k - 1$ (véase

Figura 2.3).[4]

Para probar que $d' \geq \lceil d/q \rceil$, sea $(x_{d+1}, \ldots, x_n)$ una palabra distinta de cero de $Res(C, \mathbf{c})$ y $\mathbf{x} = (x_1, \ldots, x_d, x_{d+1}, \ldots, x_n)$ su correspondiente palabra en C.

Por el principio del palomar, existe $\alpha \in F_q$ tal que al menos d/q coordenadas de $(x_1, \ldots, x_d)$ son iguales a α, por lo que se tiene que

$$d \leq wt(\mathbf{x} - \alpha\mathbf{c}) \leq d - (d/q) + wt((x_{d+1}, \ldots, x_n)).$$

pues

$$\mathbf{x} = \underbrace{x_1, \cdots}_{d/q \text{ iguales a } \alpha}, \underbrace{\ldots x_d}_{d-(d/q)\text{distintas a } \alpha}, \underbrace{x_{d+1}, \quad \cdots \quad, x_n}_{n-d \text{ distintas a } 0}.$$

Consideremos a d como si fuese el número de palomas m; y a q, como el número de palomares n. Sabemos que $d = (d/q)q$, esto es,

$$d > \left(\frac{d}{q} - 1\right) q,$$

por lo que, al menos, d/q coordenadas tienen a α.

Entonces, $d \leq d - (d/q) + d'$, esto es, $(d/q) \leq d'$ si y sólo si $\lceil d/q \rceil \leq d'$. □

2.33 Teorema (Cota de Griesmer)**.** *Sea C un código q-ario de parámetros $[n, k, d]$ donde $k \geq 1$. Entonces,*

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

Demostración. Probaremos el teorema por medio de inducción:

(i) Probamos que la desigualdad es válida para $k = 1$:

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

(ii) Cuando $k > 1$ y $\mathbf{c} \in C$ es una palabra código de peso mínimo d, el Lema 2.32 prueba que $Res(C, \mathbf{c})$ es un código $[n - d, k - 1, d']$ donde $d' \geq \lceil d/q \rceil$.

[4] Recuérdese que d debe ser minimal.

Por hipótesis inductiva, asumimos que la cota de Griesmer se satisface para $Res(C, \mathbf{c})$:

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

(iii) Para la última parte de la demostración hagamos las siguientes consideraciones: Sabemos que $\lceil d/q \rceil \leq d'$ si y sólo si $(d/q) \leq d'$. Además

$$\begin{aligned} \frac{d}{q} &\leq d' \\ \frac{d}{q^2} &\leq \frac{d'}{q} \\ &\vdots \\ \frac{d}{q^{i+1}} &\leq \frac{d'}{q^i}, \end{aligned}$$

y a su vez $(d/q^{i+1}) \leq (d'/q^i)$ si y sólo si $\lceil d/q^{i+1} \rceil \leq (d'/q^i)$. Por definición, $(d'/q^i) \leq \lceil d'/q^i \rceil$, y por transitividad

$$\left\lceil \frac{d}{q^{i+1}} \right\rceil \leq \left\lceil \frac{d'}{q^i} \right\rceil .$$

Finalmente, usando el resultado obtenido por transitividad y la hipótesis inductiva tenemos que

$$n - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^{i+1}} \right\rceil ,$$

de manera que

$$n - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^{i+1}} \right\rceil$$

y entonces

$$n \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^{i+1}} \right\rceil + d = \left\lceil \frac{d}{q} \right\rceil + \left\lceil \frac{d}{q^2} \right\rceil + \ldots + \left\lceil \frac{d}{q^{k-1}} \right\rceil + d,$$

y como

$$\left\lceil \frac{d}{q} \right\rceil + \left\lceil \frac{d}{q^2} \right\rceil + \ldots + \left\lceil \frac{d}{q^{k-1}} \right\rceil + d = \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil ,$$

el teorema queda demostrado.

n	Plotkin	Singleton	Hamming
10	2	2	2
11	2	4	3
12	2	8	5
13	2	16	7
14	2	32	11
15	4	64	16
16	4	128	26
17	6	256	40
18	10	512	64
19	20	1024	104
20	40	2048	169
21	-	4096	277
22	-	8192	460
23	-	16384	769
24	-	32768	1295
25	-	65536	2196
26	-	131072	3748

Tab. 2.1: Cotas superiores para $A_2(n, 10)$ con $10 \leq n \leq 26$

□

En la Tabla 2.1 se resumen las cotas superiores vistas a lo largo de este capítulo para $A_2(n, 10)$ donde $10 \leq n \leq 26$. Como puede apreciarse, la cota de Hamming supera por mucho a la de Singleton conforme n se hace grande. Se hace evidente también que la cota de Plotkin supera a las cotas de Singleton y a la de Hamming de forma considerable, de ahí su relevancia. Sin embargo, la cota de Plotkin no puede aplicarse a los casos en que $n > 20$, en este caso, puesto que bajo esos valores de n no se cumple que $rn < d$, lo cual se indica con un guión "-" en la tabla.

La Figura 2.4 ilustra los valores de la Tabla 2.1 con valores $15 \leq n \leq 20$.

Para finalizar el estudio del presente capítulo citaremos sólo algunos de los esfuerzos que se han hecho para mejorar, o en algunos casos encontrar, cotas para $A_q(n, d)$ durante las últimas tres décadas en orden cronológico: Best y otros (1978); Cohen y otros (1985); Janwa (1989); Conway y Sloane (1990); Tietavainen (1990); Brouwer y Verhoeff (1993); Agrell y otros (2001); Mounits y otros (2002).

Otros esfuerzos destacados que pueden encontrarse en la red son los siguientes:

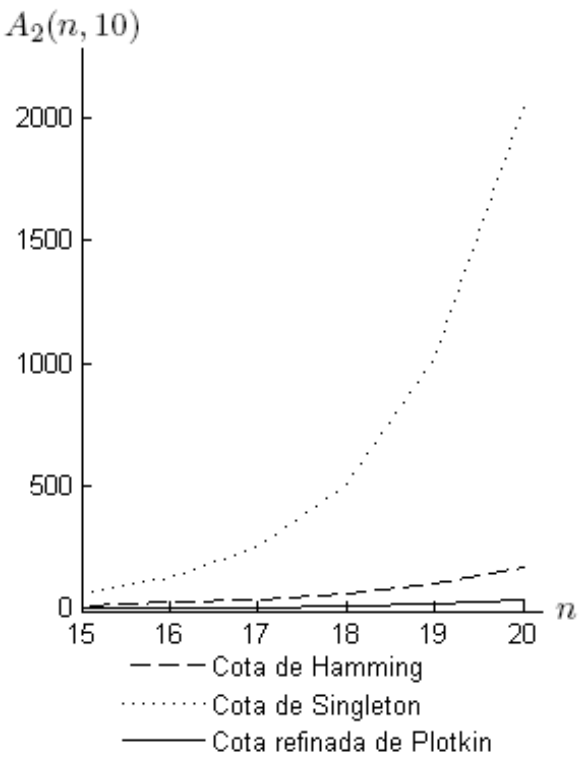


Fig. 2.4: Cotas superiores para $A_2(n, 10)$ con $15 \leq n \leq 20$

(i) el *Computer and Automation Research Institute* de la Academia Húngara de las Ciencias tiene un sitio en el que pueden descargarse en formato PDF las cotas de $A_q(n, d)$ bajo determinados parámetros. Dirección URL `http://www.sztaki.hu/~keri/codes/`, (véase Gerzson (2008));

(ii) la *Technische Universiteit Eindhoven* cuenta con un sitio al que proporciona mantenimiento Andries E. Brouwer y en el que pueden encontrarse algunas tablas de cotas para códigos q-arios con $q \leq 5$ y para códigos con pesos constantes. Dirección URL `http://www.win.tue.nl/~aeb/`, (véase Brouwer (2004));

(iii) finalmente, en la dirección `http://www.codetables.de`, Markus Grassl ofrece una aplicación en línea que permite obtener tablas de cotas para códigos lineales con parámetros específicos que el usuario debe proporcionar, (véase Grassl (2007)).

3. ANILLOS DE FROBENIUS FINITOS RELACIONADOS AL GRUPO CUÁNTICO DE LUSZTIG

"El número imaginario es un recurso bello y refinado del espíritu divino, casi un híbrido entre el ser y el no ser".
GOTTFRIED LEIBNIZ.

El artículo de Wood (1999) establece que los anillos de Frobenius finitos son los anillos más apropiados para la teoría de códigos en virtud de que dos teoremas fundamentales de MacWilliams, conocidos sobre campos finitos, se cumplen también sobre los anillos de Frobenius. Se sabe que uno de esos teoremas fundamentales, las *identidades de MacWilliams*, se cumplen en general en los códigos aditivos. El otro teorema fundamental de MacWilliams, conocido como el teorema de equivalencia de MacWilliams o la extensión del teorema de MacWilliams, se cumple sobre anillos de Frobenius finitos. La extensión del teorema de MacWilliams aborda la noción de equivalencia de códigos. Dos códigos son equivalentes si existe una transformación monomial que lleve de un código a otro. Si dos códigos lineales son isomórficos como espacios vectoriales abstractos y existe un isomorfismo que preserve el peso de Hamming, entonces dicho isomorfismo se extiende a una transformación monomial.

Un artículo reciente de Wood (2006) muestra que el resultado a la inversa también es válido: si la extensión del teorema de MacWilliams se cumple para un anillo finito, entonces el anillo debe ser de Frobenius. Esto es una fuerte evidencia de que los anillos de Frobenius son los anillos más apropiados para la teoría de códigos.

Por estas razones el problema de la construcción de anillos de Frobenius finitos ha cobrado vital importancia. Afortunadamente, hay una rama del álgebra moderna - la teoría de los grupos cuánticos - que proporciona ejemplos de álgebras de Frobenius. El teorema de Larson y Sweedler (1969) prueba que toda álgebra de Hopf de dimensión finita en realidad es de Frobenius, mientras que un estudio reciente de Skryabin (2007) señala que cada subálgebra de coidealo de un sólo lado también es un álgebra de Frobenius. Estos teoremas proporcionan ejemplos importantes de anillos de Frobenius finitos si consideramos álgebras de Hopf y sus subálgebras con coidealo derecho sobre un campo finito F_q con $q = p^n$ elementos (donde p es un número primo).

En este capítulo, consideraremos códigos sobre anillos de Frobenius finitos relacionados al grupo cuántico pequeño de Lusztig.

A lo largo de la segunda mitad de la presente tesis el concepto de caracter es fundamental. Para comprender cabalmente este concepto haremos un breve repaso de la forma polar de los números complejos.

3.1. La forma polar de los números complejos

Los números complejos son arreglos o vectores bidimensionales por lo que resulta natural identificarlos como puntos del plano $\mathbb{R}^2$ para obtener una representación geométrica del sistema de los números complejos. Haciendo la consideración anterior, podemos referirnos entonces a un punto con coordenadas específicas como un número complejo y al plano $\mathbb{R}^2$ como el plano complejo.

3.1 Definición. El *valor absoluto* o *módulo* de un número complejo $z = (x, y)$ es la longitud del vector (x, y), es decir,

$$|z| = \sqrt{x^2 + y^2}.$$

Puesto que el valor absoluto de un número complejo z es la longitud de un vector, tiene las siguientes propiedades:

(i) $|z| \geq 0$; $|z| = 0$ si y sólo si $z = 0$,

(ii) $|rz| = r\,|z|$ para $r \in \mathbb{R}$ y $z \in \mathbb{C}$,

(iii) $|z_1 + z_2| \leq |z_1| + |z_2|$ (desigualdad del triángulo).

Se sabe que

$$z\overline{z} = x^2 + y^2 = |z|^2 .$$

En el plano complejo, el punto $\overline{z}$ es la imagen simétrica, respecto al eje X, del punto z.

El módulo de los números complejos tiene además las siguientes propiedades adicionales a las citadas. Desarrollando

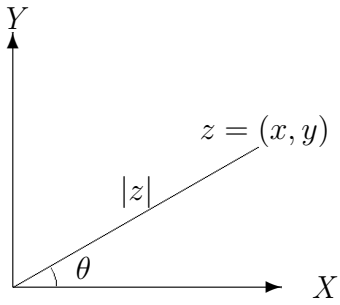


Fig. 3.1: Representación de los complejos en el plano $\mathbb{R}^2$

$$\begin{aligned} |z_1 z_2|^2 &= (z_1 z_2)\overline{(z_1 z_2)} \\ &= (z_1 z_2)(\overline{z}_1 \overline{z}_2) \\ &= (z_1 \overline{z_2})(z_1 \overline{z_2}) \\ &= |z_1|^2 |z_2|^2 , \end{aligned}$$

esto es,

$$|z_1 z_2| = |z_1| \, |z_2| . \tag{3.1}$$

Además, como

$$|z_1| = \left| \frac{z_1}{z_2} z_2 \right| = \left| \frac{z_1}{z_2} \right| |z_2| ,$$

y dividiendo por $|z_2|$ se tiene que

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|} \qquad \text{siempre que } z_2 \neq 0. \tag{3.2}$$

3.2 Definición. El *argumento* o *amplitud* de un número complejo distinto de cero $x = (x, y)$, denotado como $Arg(z)$ o $Am(z)$, es una medida en radianes del ángulo de inclinación del vector (x, y).

En la Figura 3.1 se aprecia la representación geométrica de los números complejos.

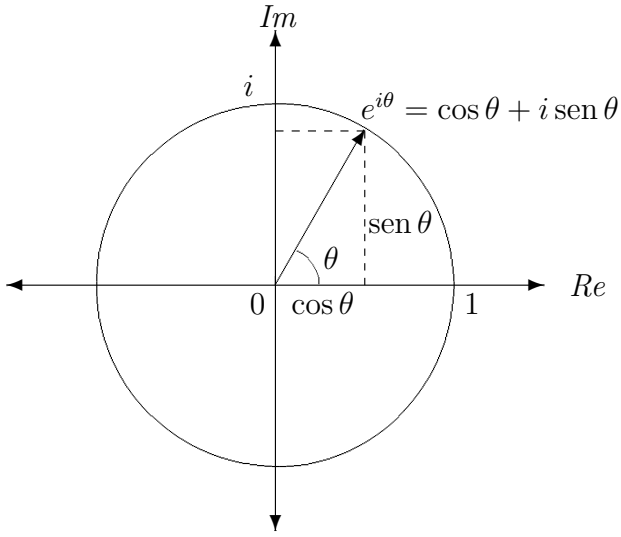


Fig. 3.2: La exponencial compleja de Euler.

$\mathbf{P} = (x, y)$ puede representarse por medio de coordenadas polares (r, θ) donde

$$\begin{cases} x = r\cos\theta, \\ y = r\mathrm{sen}\,\theta, \end{cases}$$

por lo que el número complejo $z = x + iy$ puede escribirse, en su forma polar, también como

$$z = r(\cos\theta) + i\,\mathrm{sen}\,\theta$$

en donde $r = |z|$ y $\theta = Arg(z)$.

Introduzcamos ahora la *exponencial compleja* $e^{i\theta}$, definida por la relación

$$e^{i\theta} = \cos\theta + i\,\mathrm{sen}\,\theta, \tag{3.3}$$

cuya representación geométrica se aprecia en la Figura 3.2.

La forma polar del número complejo $z = x + iy$ puede expresarse, en términos de la exponencial compleja, como

$$z = re^{i\theta}.$$

Esta forma polar es particularmente útil para simplificar las operaciones de multiplicación y división, así como para calcular las potencias y las raíces de los números complejos.

Empleando la exponencial compleja (3.3), y puesto que $\cos(a+b) = \cos a \cos b - \operatorname{sen} a \operatorname{sen} b$ y $\operatorname{sen}(a+b) = \cos a \operatorname{sen} b + \cos b \operatorname{sen} a$, tenemos que

$$\begin{aligned} e^{i\theta_1} e^{i\theta_2} &= (\cos\theta_1 + i \operatorname{sen}\theta_1)(\cos\theta_2 + i \operatorname{sen}\theta_2) \\ &= (\cos\theta_1 \cos\theta_2 - \operatorname{sen}\theta_1 \operatorname{sen}\theta_2) + i(\cos\theta_1 \operatorname{sen}\theta_2 + \operatorname{sen}\theta_1 \cos\theta_2) \\ &= \cos(\theta_1+\theta_2) + i \operatorname{sen}(\theta_1+\theta_2) \\ &= e^{i(\theta_1+\theta_2)}. \end{aligned} \tag{3.4}$$

Por otra parte,

$$\frac{1}{e^{i\theta}} = e^{-i\theta},$$

puesto que, por (3.4),

$$e^{i\theta} e^{-i\theta} = e^{i0} = (\cos 0, \operatorname{sen} 0) = 1.$$

Por lo tanto

$$\frac{e^{i\theta_1}}{e^{i\theta_2}} = e^{i\theta_1} e^{-i\theta_2} = e^{i(\theta_1-\theta_2)}. \tag{3.5}$$

Esto es, para multiplicar exponenciales complejas sumamos los exponentes y, para dividirlas, restamos los exponentes. Empleando las ecuaciones (3.1), (3.2), (3.4) y (3.5), tenemos

$$z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1+\theta_2)} \tag{3.6}$$

y si $z_2 \neq 0$, entonces

$$\frac{z_1}{z_2} = \frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}. \tag{3.7}$$

Esto significa que para multiplicar dos números complejos en forma polar multiplicamos los módulos y sumamos los argumentos y, para dividir dos números complejos en forma polar, dividimos los módulos y restamos los argumentos. Es decir,

$$|z_1 z_2| = |z_1|\,|z_2|\,, \qquad Arg(z_1 z_2) = Arg(z_1) + Arg(z_2),$$

y para $z_2 \neq 0$

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \qquad Arg\left(\frac{z_1}{z_2}\right) = Arg(z_1) - Arg(z_2).$$

Usando la ecuación (3.6) e inducción matemática puede mostrarse que

$$|z_1 z_2 \cdots z_n| = |z_1|\,|z_2| \cdots |z_n|$$

y que

$$Arg(z_1 z_2 \cdots z_n) = Arg(z_1) + Arg(z_2) + \ldots + Arg(z_n).$$

Si en estas igualdades tomamos todos los n complejos iguales, tenemos que

$$|z^n| = |z|^n = r^n$$

y

$$Arg(z^n) = nArg(z) = n\theta,$$

de donde obtenemos la fórmula conocida como *fórmula de De Moivre*:

$$z^n = (re^{i\theta})^n = r^n e^{in\theta}. \tag{3.8}$$

Consideremos ahora el problema de encontrar las raíces n-ésimas de un número complejo z. Por una raíz n-ésima de z entendemos un número complejo cualquiera ζ tal que $\zeta^n = z$. Sea

$$\zeta = \sigma e^{i\varphi}$$

una raíz n-ésima de $z = re^{i\theta}$. Es decir

$$\zeta^n = \sigma^n e^{in\varphi} = re^{i\theta}.$$

Esta última ecuación implica que $\sigma^n = r$ y que $e^{in\varphi} = e^{i\theta}$. Si $\sigma^n = r$, entonces $\sigma = r^{1/n}$ donde $r^{1/n}$ es el número real no negativo cuya potencia n-ésima es r. Ahora, para los números θ_1 y θ_2, se sigue que $e^{i\theta_1} = e^{i\theta_1}$ si y sólo si

$$e^{i(\theta_1 - \theta_2)} = 1,$$

lo cual implica que $\cos(\theta_1 - \theta_2) = 1$ y $\operatorname{sen}(\theta_1 - \theta_2) = 0$, y esto sucede siempre que

$$\theta_1 - \theta_2 = 2\pi k \qquad \text{para } k \in \mathbb{Z}.$$

De aquí que $e^{in\varphi} = e^{i\theta}$ si y sólo si $n\varphi = \theta + 2\pi k$ para un k entero o

$$\varphi = \frac{\theta + 2\pi k}{n}.$$

Entre los valores antes dados para φ, cualesquiera que difieran en un múltiplo entero de 2π resultarán naturalmente en el mismo número complejo ζ. Por esta razón resulta suficiente considerar sólo $k \in [0, n-1]$ para obtener todas las raíces n-ésimas distintas a z. Entonces, si

$$z = re^{i\theta},$$

para un entero positivo n, los números

$$r^{1/n} e^{i(\theta+2\pi k)/n}$$

donde $k = 0, 1, 2, \ldots, n-1$ son las raíces n-ésimas de z.

Las raíces n-ésimas de z se encuentran sobre una circunferencia de radio $r^{1/n}$ con centro en el origen y se encuentran igualmente espaciadas, cada una de ellas con un argumento igual a $1/nArg(z)$.

3.3 Ejemplo. Encontremos todas las raíces cuartas de $z = 16 - 16\sqrt{3}i$. El valor absoluto de z está dado por $r = |z| = 32$. Sabemos que $\theta_1 = \pi/3$ satisface

$$\cos\theta = 1/2,$$
$$\operatorname{sen}\theta = \sqrt{3}/2$$

y entonces, $Arg(z) = 2\pi - \pi/3 = 5\pi/3$ (Figura 3.1). Como $n = 4$, las raíces de z tienen la forma

$$r^{1/n} e^{i(\theta+2\pi k)/n} = \sqrt[4]{32} e^{i(5\pi/12 + \pi k/2)} \qquad k = 0, 1, 2, 3.$$

Entonces las raíces cuartas de z son:

$$\zeta_1 = \sqrt[4]{32} e^{i(5\pi/12)}$$
$$\zeta_2 = \sqrt[4]{32} e^{i(11\pi/12)}$$
$$\zeta_3 = \sqrt[4]{32} e^{i(17\pi/12)}$$
$$\zeta_4 = \sqrt[4]{32} e^{i(23\pi/12)}.$$

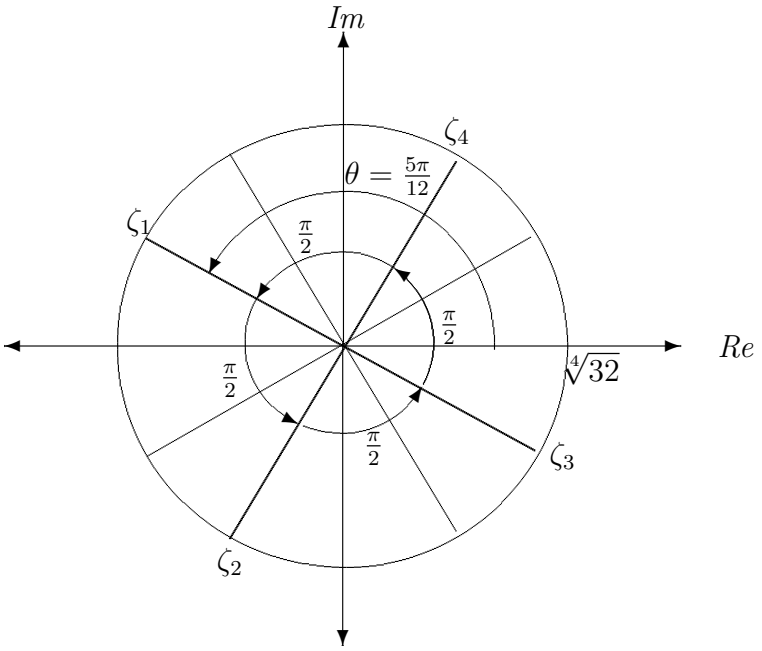


Fig. 3.3: Raíces cuartas de $z = 16 - 16\sqrt{3}i$.

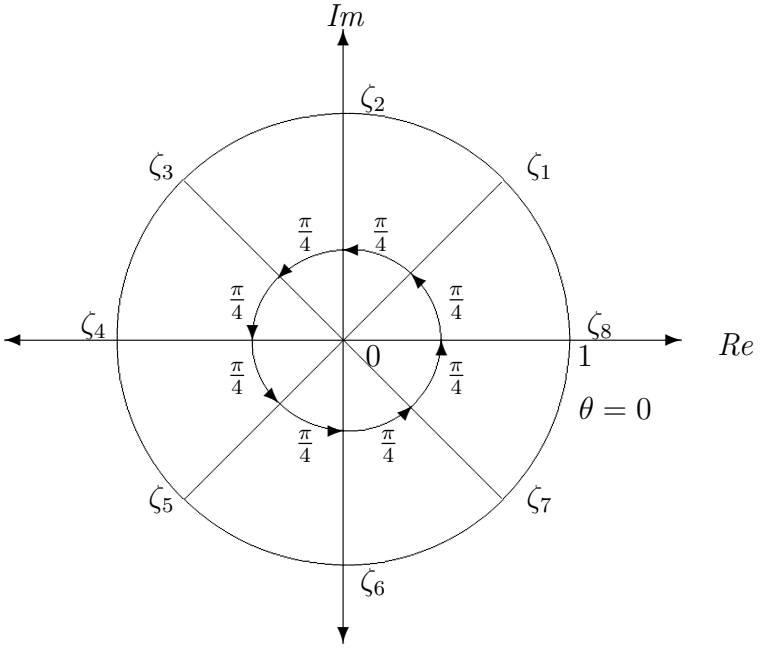


Fig. 3.4: Raíces octavas de $z = 1$.

θ	0	$\pi/4$	$\pi/2$	$3\pi/4$	π	$5\pi/4$	$3\pi/2$	$7\pi/4$
$x = \cos\theta$	1	$\sqrt{2}/2$	0	$-\sqrt{2}/2$	-1	$-\sqrt{2}/2$	0	$\sqrt{2}/2$
$y = \operatorname{sen}\theta$	0	$\sqrt{2}/2$	1	$\sqrt{2}/2$	0	$-\sqrt{2}/2$	-1	$-\sqrt{2}/2$

Tab. 3.1: Valores del seno y coseno en radianes dados.

3.4 Ejemplo. Ahora, calcularemos todas las raíces octavas de $z = 1$, esto es todas las ζ tales que $\zeta^8 = 1$. Naturalmente, $|z| = 1$ y $Arg(z) = 0 = \theta$.

Como $n = 8$, las raíces de z tienen la forma

$$r^{1/n}e^{i(\theta+2\pi k)/n} = e^{i(\pi k/4)} \qquad k = 0, 1, 2, 3.$$

Entonces las raíces cuartas de z son:

$$\zeta_1 = e^0, \quad \zeta_2 = e^{i(\pi/4)}, \quad \zeta_3 = e^{i(\pi/2)}, \quad \zeta_4 = e^{i(3\pi/4)}, \quad \zeta_5 = e^{i(\pi)},$$
$$\zeta_6 = e^{i(5\pi/4)}, \quad \zeta_7 = e^{i(6\pi/4)}, \quad \zeta_8 = e^{i(7\pi/4)},$$

mismas que pueden apreciarse en la Figura 3.4.

3.5 Definición. Se dice que una álgebra A es un *álgebra de Frobenius* si existe una forma bilineal asociativa no degenerada $(\ ,\ ) : A \otimes A \to F_q$.

Recordemos que una forma es bilineal si y sólo si

$$(\alpha a + \beta b, c) = \alpha(a, c) + \beta(b, c), \quad (a, \beta b + \gamma c) = \beta(a, b) + \gamma(a, c)$$

para toda $a, b, c \in A$. Es asociativa si

$$(a, bc) = (ab, c)$$

para toda $a, b, c \in A$. Es no degenerada si $(A, b) = 0$ implica $b = 0$, y $(a, A) = 0$ implica $a = 0$.

3.2. Construcción del caracter principal de un álgebra de Frobenius finita

Denotemos como ζ a la p-ésima raíz primitiva de 1 en el campo de los números complejos $\mathbb{C}$. Todo elemento $\alpha \in F_p$ tiene una representación $\alpha = \underbrace{1 + 1 + \cdots + 1}_{m}$ (recordemos

que el campo F_p es isomórfico a $\mathbb{Z}/p\mathbb{Z}$). Primero, definimos $\chi_0(\alpha) = \zeta^m$. Entonces, el caracter principal χ para un álgebra R de dimensión finita definida sobre F_p se define como sigue:

$$\chi(a) = \chi_0(\alpha), \tag{3.9}$$

donde $\alpha = (a, 1) = (1, a) \in F_p$.

3.6 Definición. El *caracter* de un anillo finito R es un mapeo $\chi : R \to \mathbb{C}$ de R al campo de los números complejos $\mathbb{C}$ que satisface

$$\chi(a + b) = \chi(a)\chi(b). \tag{3.10}$$

3.7 Ejemplo. De acuerdo con el Ejemplo 3.4, y considerando los valores de la Tabla 3.1, los caracteres de los elementos de $\mathbb{Z}_8$ son:

$$\chi(0) = 1, \quad \chi(1) = \sqrt{2}/2 + i\sqrt{2}/2, \quad \chi(2) = i, \quad \chi(3) = -\sqrt{2}/2 + i\sqrt{2}/2,$$

$$\chi(4) = -1, \quad \chi(5) = -\sqrt{2}/2 - i\sqrt{2}/2, \quad \chi(6) = -i, \quad \chi(7) = \sqrt{2}/2 - i\sqrt{2}/2.$$

Note que la propiedad (3.10) se verifica para cualquier pareja de caracteres que se elija.

3.8 Definición. El caracter χ se llama *caracter principal derecho* si para cada caracter distinto χ_1 existe un elemento $a \in R$ tal que $\chi_1(x) = \chi(ax)$. De la misma forma, se dice que χ es un *caracter principal izquierdo* si para cada caracter distinto χ_1 existe un elemento $b \in R$ tal que $\chi_1(x) = \chi(xb)$.

3.9 Ejemplo. Identifiquemos los caracteres principales del ejemplo anterior. Se observa que:

$$\chi_2(x) = \chi_1(2x), \quad \chi_3(x) = \chi_1(3x), \quad \chi_4(x) = \chi_1(4x), \quad \chi_5(x) = \chi_1(5x),$$

$$\chi_6(x) = \chi_1(6x), \quad \chi_7(x) = \chi_1(7x), \quad \chi_0(x) = \chi_1(0x).$$

Esto es, $\chi_1(x)$ es caracter principal. Lo mismo se observa para $\chi_3(x)$,

$$\chi_6(x) = \chi_3(2x), \quad \chi_1(x) = \chi_3(3x), \quad \chi_4(x) = \chi_3(4x), \quad \chi_7(x) = \chi_3(5x),$$

$$\chi_2(x) = \chi_3(6x), \quad \chi_5(x) = \chi_3(7x), \quad \chi_0(x) = \chi_3(0x);$$

para $\chi_5(x)$,

$$\chi_2(x) = \chi_5(2x), \quad \chi_7(x) = \chi_5(3x), \quad \chi_4(x) = \chi_5(4x), \quad \chi_1(x) = \chi_5(5x),$$

$$\chi_6(x) = \chi_5(6x), \quad \chi_3(x) = \chi_5(7x), \quad \chi_0(x) = \chi_5(0x);$$

y para $\chi_7(x)$,

$$\chi_6(x) = \chi_7(2x), \quad \chi_5(x) = \chi_7(3x), \quad \chi_4(x) = \chi_7(4x), \quad \chi_3(x) = \chi_7(5x),$$

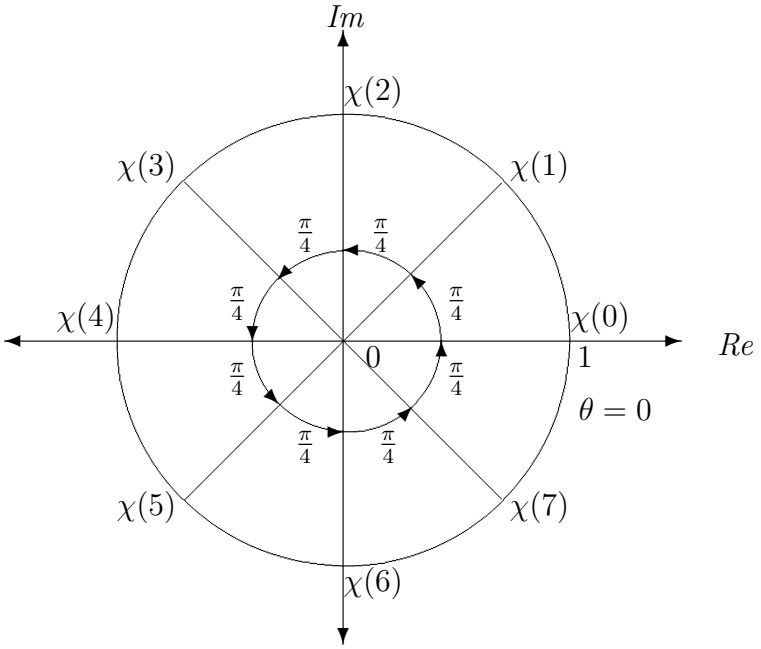


Fig. 3.5: Caracteres de $\mathbb{Z}_8$.

$$\chi_2(x) = \chi_7(6x), \quad \chi_1(x) = \chi_7(7x), \quad \chi_0(x) = \chi_7(0x), ;$$

por lo que se tiene que tanto $\chi_1(x)$, $\chi_3(x)$, $\chi_5(x)$ y $\chi_7(x)$ son caracteres principales de $\mathbb{Z}_8$.

El resto de los caracteres tienen un patrón cíclico que impide que a través de sí mismos se formen todos los demás caracteres. Por ejemplo, en el caso de $\chi_2(x)$:

$$\chi_4(x) = \chi_2(2x), \quad \chi_6(x) = \chi_2(3x), \quad \chi_0(x) = \chi_2(4x), \quad \chi_2(x) = \chi_2(5x),$$

$$\chi_4(x) = \chi_2(6x), \quad \chi_6(x) = \chi_2(7x), \quad \chi_0(x) = \chi_2(0x).$$

Y de forma similar, para $\chi_4(x)$

$$\chi_0(x) = \chi_4(2x), \quad \chi_4(x) = \chi_4(3x), \quad \chi_0(x) = \chi_4(4x), \quad \chi_4(x) = \chi_4(5x),$$

$$\chi_0(x) = \chi_4(6x), \quad \chi_4(x) = \chi_4(7x), \quad \chi_0(x) = \chi_4(0x);$$

para $\chi_6(x)$

$$\chi_4(x) = \chi_6(2x), \quad \chi_2(x) = \chi_6(3x), \quad \chi_0(x) = \chi_6(4x), \quad \chi_6(x) = \chi_6(5x),$$

$$\chi_4(x) = \chi_6(6x), \quad \chi_2(x) = \chi_6(7x), \quad \chi_0(x) = \chi_6(0x);$$

y para $\chi_0(x)$

$$\chi_0(x) = \chi_0(2x) = \chi_0(3x) = \chi_0(4x) = \chi_0(5x),$$

$$\chi_0(x) = \chi_0(6x) = \chi_0(7x) = \chi_0(0x);$$

por lo que se tiene que tanto $\chi_2(x)$, $\chi_4(x)$, $\chi_6(x)$ y $\chi_0(x)$ *no* son caracteres principales de $\mathbb{Z}_8$.

Una de las caracterizaciones más importantes de los anillos de Frobenius finitos dice que un anillo finito R es de Frobenius si y sólo si R posee un caracter principal derecho (o, de forma equivalente, un caracter principal izquierdo), véase Honold (2001).

3.10 Ejemplo. Del Ejemplo 3.9 se tiene que $\mathbb{Z}_8$ cuenta con cuatro caracteres principales, esto es, $\mathbb{Z}_8$ es un anillo de Frobenius.[1]

3.3. Algoritmo de reducción en la forma de Lusztig

Consideremos el conjunto de variables no conmutativas $x_1, x_2, \ldots, x_n$. Fijemos un parámetro de cuantificación q en el campo finito F_p de p elementos (donde p es un número primo) tal que $q^4 \neq 1$. Por medio de este parámetro construiremos una forma multiplicativa P sobre el conjunto de todas las palabras en $x_1, x_2, \ldots, x_n$ como se explica a continuación. Primero, definimos los generadores de esta forma:

$$P(x_i, x_j) = \begin{cases} q^{-1} & \text{si } i = j \pm 1 \ ; \\ q^2 & \text{si } i = j; \\ 1 & \text{si } |i - j| > 1. \end{cases}$$

Entonces, para un par arbitrario de palabras u, v el valor de $P(u, v)$ puede encontrarse por medio de la bimultiplicatividad:

$$P(u, vw) = P(u, v)P(u, w), \quad P(uv, w) = P(u, w)P(v, w).$$

3.11 Ejemplo. Si $u = x_1x_2$, $v = x_2x_3$, entonces tenemos

[1] Puesto que $\mathbb{Z}_8$ es un anillo conmutativo, los caracteres principales izquierdos son iguales a los caracteres principales derechos.

$$\begin{aligned} P(u,v) &= P(x_1x_2, v) \\ &= P(x_1, v)P(x_2, v) \\ &= P(x_1, x_2x_3) \cdot P(x_2, x_2x_3) \\ &= P(x_1, x_2)P(x_1, x_3) \cdot P(x_2, x_2)P(x_2, x_3) \\ &= q^{-1} \cdot 1 \cdot q^2 q^{-1} \\ &= 1. \end{aligned}$$

3.12 Ejemplo. Sea $u = x_1x_2 = v$. Tenemos entonces

$$\begin{aligned} P(u,v) &= P(x_1x_2, x_1x_2) \\ &= P(x_1, x_1x_2) \cdot P(x_2, x_1x_2) \\ &= P(x_1, x_1)P(x_1, x_2) \cdot P(x_2, x_1)P(x_2, x_2) \\ &= q^2q^{-1} \cdot q^{-1}q^2 \\ &= q^2. \end{aligned}$$

3.13 Definición. Con la ayuda de la forma P, se define una nueva operación, *corchetes torcidos*, como sigue:

$$[u, v] = uv - P(u, v)vu.$$

3.14 Ejemplo. $[x_1, x_2] = x_1x_2 - q^{-1}x_2x_1$. Al mismo tiempo $[x_2, x_1] = x_2x_1 - q^{-1}x_1x_2$.

Algunos otros ejemplos se siguen de los ejemplos considerados con anterioridad para calcular P :

$$[x_1x_2, x_2x_3] = x_1x_2^2x_3 - x_2x_3x_1x_2;$$

$$[x_1x_2, x_1x_2] = x_1x_2 \cdot x_1x_2 - q^2x_1x_2 \cdot x_1x_2 = (1 - q^2)(x_1x_2)^2.$$

Esta operación está definida para todos los pares de palabras en $x_1, x_2, \ldots, x_n$. Ahora, extenderemos la definición de corchetes torcidos a las combinaciones lineales de palabras por medio de las siguientes condiciones lineales:

$$[\alpha u + \beta v, w] = \alpha[u, w] + \beta[v, w],$$

$$[u, \beta v + \gamma w] = \beta[u, v] + \gamma[u, w].$$

De esta forma podremos realizar cálculos más complicados con el operador corchetes torcidos tal como en el siguiente ejemplo.

3.15 Ejemplo. (i) Sabemos que $[x_1, x_2] = x_1x_2 - q^{-1}x_2x_1$, y que

$$P(x_1x_2, x_2) = P(x_1, x_2)P(x_2, x_2) = q^{-1}q^2 = q = P(x_2x_1, x_2).$$

Entonces,

$$\begin{aligned}
[[x_1, x_2], x_2] &= [x_1x_2 - q^{-1}x_2x_1, x_2] \\
&= [x_1x_2, x_2] - q^{-1}[x_2x_1, x_2] \\
&= x_1x_2^2 - P(x_1x_2, x_2)x_2x_1x_2 - q^{-1}(x_2x_1x_2 - P(x_2x_1, x_2)x_2x_2x_1) \\
&= x_1x_2^2 - qx_2x_1x_2 - q^{-1}x_2x_1x_2 + x_2x_2x_1 \\
&= x_1x_2^2 - (q + q^{-1})x_2x_1x_2 + x_2^2x_1.
\end{aligned}$$

(ii) Sabemos que $P(x_1x_2, x_3) = P(x_1, x_3)P(x_2, x_3) = 1q^{-1} = q^{-1} = P(x_2x_1, x_3)$. Entonces,

$$\begin{aligned}
[[x_1, x_2], x_3] &= [x_1x_2 - q^{-1}x_2x_1, x_3] \\
&= [x_1x_2, x_3] - q^{-1}[x_2x_1, x_3] \\
&= x_1x_2x_3 - P(x_1x_2, x_3)x_3x_1x_2 - q^{-1}(x_2x_1x_3 - P(x_2x_1, x_3)x_3x_2x_1) \\
&= x_1x_2x_3 - q^{-1}x_3x_1x_2 - q^{-1}x_2x_1x_3 + q^{-2}x_3x_2x_1.
\end{aligned}$$

3.16 Definición. El *álgebra cuántica de Borel* U_q^+ (de tipo A_n) se define a través de las siguientes relaciones, conocidas como las *relaciones cuánticas de Serre*:

$$[x_i, x_j] = 0, \quad \text{si } |i - j| > 1; \tag{3.11}$$

$$[[x_i, x_{i+1}], x_{i+1}] = 0, \quad \text{si } 1 \leq i < n; \tag{3.12}$$

$$[x_i, [x_i, x_{i+1}]] = 0, \quad \text{si } 1 \leq i < n. \tag{3.13}$$

El álgebra U_q^+ no es de dimensión finita. Sin embargo, tiene un conjunto finito de generadores PBW.[2]

3.17 Definición. Un subconjunto linealmente ordenado $W \subseteq U_q^+$ se dice un *conjunto de generadores PBW* de U_q^+ si existe una función $h : W \to \mathbf{Z}^+ \cup \infty$, conocida como la *función altura*, tal que el conjunto de todos los productos

$$w_1^{n_1}w_2^{n_2}\cdots w_k^{n_k}, \tag{3.14}$$

donde $j \in J$, $w_1 < w_2 < \ldots < w_k \in W$, $n_i < h(w_i), 1 \leq i \leq k$ es una base de U_q^+. Al valor $h(w)$ se le llama *altura* de w en W.

[2] PBW por Poincarè-Birkhoff-Witt, quienes construyeron este tipo de bases para las álgebras envolventes universales asociativas de las álgebras de Lie.

Para encontrar el conjunto W de generadores PBW para U_q^+, consideremos los siguientes polinomios

$$w_{k,m} \stackrel{df}{=} [\dots[[x_k, x_{k+1}], x_{k+2}], \dots, x_m], \tag{3.15}$$

donde $1 \leq k \leq m \leq n$. Evidentemente, el número de elementos $w_{k,m}$ es igual al número de pares (k, m), $k \leq m$. Lo cual suma un total de $(n(n-1)/2) + n = n(n+1)/2$.

El conjunto W es un conjunto de generadores PBW para U_q^+ dado que definimos el orden en W como sigue:

$$w_{k,m} > w_{s,t} \text{ si } k < s \text{ o } k = s, m < t.$$

Dicho de otro modo, todo elemento $a \in U_q^+$ tiene una representación única como una combinación lineal de productos

$$a = \sum \alpha_i w_{k_{1i},m_{1i}}^{n_{1i}} w_{k_{2i},m_{2i}}^{n_{2i}} \dots w_{k_{ri},m_{ri}}^{n_{ri}}, \tag{3.16}$$

donde $w_{k_{1i},m_{1i}} < w_{k_{2i},m_{2i}} < \dots < w_{k_{ri},m_{ri}}$ con respecto al orden definido previamente para el conjunto W.

3.18 Ejemplo. Si $n = 3$ entonces existen sólo $3(3+1)/2 = 6$ elementos $w_{k,m}$:

$$w_{1,1} = x_1;\ w_{1,2} = [x_1, x_2];\ w_{1,3} = [[x_1, x_2], x_3];\ w_{2,2} = x_2;\ w_{2,3} = [x_2, x_3];\ w_{3,3} = x_3.$$

De acuerdo con la definición anterior estos elementos están ordenados como sigue:

$$x_3 < [x_2, x_3] < x_2 < [[x_1, x_2], x_3] < [x_1, x_2] < x_1.$$

Por lo tanto, por definición de los generadores PBW, cada elemento del álgebra U_q^+ con $n = 3$ es una combinación lineal de elementos linealmente independientes:

$$x_3^{n_1}[x_2, x_3]^{n_2} x_2^{n_3}[[x_1, x_2], x_3]^{n_4}[x_1, x_2]^{n_5} x_1^{n_6}. \tag{3.17}$$

Resulta entonces que todas las alturas de los elementos $w_{k,m}$ en U_q^+ son infinitas. En particular, la base (3.17) es infinita. Entonces U_q^+ es un álgebra infinita. Para hacerla finita, deberemos aplicar la modificación de Lusztig. Sea N el orden multiplicativo del parámetro de cuantificación q^2 sobre el campo F_p; esto es, $q^{2N} = 1$, mientras que $q^{2m} \neq 1$, $m < N$.

La modificación de Lusztig aparece de U_q^+ por medio de las siguientes relaciones adicionales

$$w_{k,m}^N = 0, \quad 1 \leq k \leq m \leq n. \tag{3.18}$$

Denotaremos esta nueva álgebra sobre F_p como u_q^+. Las relaciones (3.18) muestran que todos los generadores PBW $w_{k,m}$ tienen la misma altura finita en u_q^+, misma que es igual a N. En particular, la dimensión de u_q^+ es igual al número M de posibles elementos

$$w_{k_1,m_1}^{n_1} w_{k_2,m_2}^{n_2} \dots w_{k_s,m_s}^{n_s}, \quad n_i < N, \tag{3.19}$$

donde $s = n(n+1)/2$ es el número total de pares (k, m), $1 \leq 1 \leq k \leq m \leq n$, y

$$w_{k_1,m_1} < w_{k_2,m_2} < \ldots < w_{k_s,m_s}.$$

Así, el número M es igual a $N^{n(n+1)/2}$. Puesto que el campo F_p tiene p elementos, el número total de elementos en u_q^+ es igual a $p^{N^{n(n+1)/2}}$. Particularmente, u_q^+ es un anillo finito.

Además u_q^+ es un álgebra de Frobenius sobre F_p. Para comprobarlo, deberemos definir una forma bilineal asociativa no degenerada. Denotemos como U el elemento menor posible como elemento base (3.19) con respecto al orden lexicográfico de palabras en $w_{k,m}$. El elemento U tiene la forma

$$U = x_n^{N-1}[x_{n-1}, x_n]^{N-1} x_{n-1}^{N-1}[[x_{n-2}, x_{n-1}], x_n]^{N-1} \ldots.$$

Si a, b son dos elementos de u_q^+, entonces el elemento ab es una combinación lineal de los elementos base (3.19). En particular, el elemento U toma parte en esta combinación con un coeficiente único denotado como $\alpha \in F_p$; esto es, $ab = \alpha U + \cdots$, donde los puntos suspensivos denotan una combinación lineal de los elementos base (3.19) diferente de U. Definimos

$$(a, b) = \alpha.$$

Esta forma es asociativa, para (a, bc) y (ab, c) que están definidas por la descomposición de los mismos elementos $abc = \beta U + \cdots$; esto es, $(a, bc) = (ab, c) = \beta$, lo cual es una forma bilineal y no degenerada.

Ahora describiremos un algoritmo que permite encontrar una representación de la forma de (3.19) para cualquier producto de elementos base. Esto, a su vez, nos permitirá calcular la forma y por lo tanto el caracter principal (véase (3.9)). Se sabe que las relaciones (3.12-3.13) implican las siguientes relaciones entre los generadores PBW $w_{k,m}$.

$$[w_{k,m}, w_{s,t}] = \begin{cases} w_{k,t}, & \text{si } s = m+1; \\ 0, & \text{si } s \neq m+1, k \leq s, \text{ y } (k,m) \neq (s,t). \end{cases} \tag{3.20}$$

El miembro izquierdo de esta relación es igual a $w_{k,m}w_{s,t} - \alpha w_{s,t}w_{k,m}$, donde el coeficiente α se define con la ayuda de la forma P como sigue

$$\alpha = P(x_k x_{k+1} \ldots x_m, x_s x_{s+1} \ldots x_t) = \prod_{k \leq i \leq m} \prod_{s \leq j \leq t} P(x_i, x_j).$$

Ahora, consideraremos el conjunto W de generadores PBW como un nuevo alfabeto. Sabemos que todo elemento $a \in u_q^+$ tiene una representación única (3.16) como una combinación lineal de palabras monótonas en W. Consideremos el orden lexicográfico del conjunto de todas las palabras en W (no sólo las monótonas). La propiedad más importante de las relaciones (3.20) es que permiten reemplazar la palabra $w_{k,m}w_{s,t}$ con

$w_{k,m} > w_{s,t}$ por la suma de menos palabras en W. De manera más precisa, ésta es la suma de $w_{s,t}w_{k,m}$ y $w_{s,m}$ multiplicada por algún coeficiente (posiblemente cero). De esta forma se tiene un algoritmo para reducir un producto arbitrario, y no necesariamente monótono, de elementos de W a una combinación lineal de palabras monótonas (y por lo tanto palabras base) en W.

Sea a una palabra en W. Si a no tiene subpalabras $w_{k,m}w_{s,t}$ tales que $w_{k,m} > w_{s,t}$, entonces es monótona y ya no hay nada que hacer. Si existe una subpalabra $w_{k,m}w_{s,t}$ tal que $w_{k,m} > w_{s,t}$, entonces aplicamos la relación (3.20). Como resultado, obtendremos la suma de una o dos palabras menores. Luego continuamos este proceso tanto como sea posible. El algoritmo se detendrá cuando las palabras en la combinación lineal resultante carezcan de subpalabras de la forma $w_{k,m}w_{s,t}$ tales que $w_{k,m} > w_{s,t}$; esto es, cuando todas ellas sean monótonas, que es lo que se persigue.

Consideremos algunos ejemplos.

3.19 Ejemplo. (i) Supongamos que $n = 2$; esto es, tenemos sólo los dos generadores x_1, x_2. Respectivamente, el conjunto W de generadores PBW tiene $2(2+1)/2 = 3$ elementos

$$w_{2,2} = x_2,\ w_{1,2} = [x_1, x_2],\ w_{1,1} = x_1.$$

En particular, cada elemento $a \in U_q^+$ (con $n = 2$) tiene una representación única

$$a = \sum_i \alpha_i x_2^{n_{1i}} [x_1, x_2]^{n_{2i}} x_1^{n_{3i}},$$

mientras que la palabra minimal (en el sentido del orden lexicográfico de palabras en W) es

$$U = x_2^{N-1}[x_1, x_2]^{N-1}x_1^{N-1}.$$

Las relaciones (3.20) toman la forma

$$[w_{1,2}, w_{2,2}] = [[x_1, x_2], x_2] = 0,\ [w_{1,1}, w_{1,2}] = [x_1, [x_1, x_2]] = 0,\ [w_{1,1}, w_{2,2}] = w_{1,2}. \tag{3.21}$$

Sea $a = x_1$, $b = x_2$. Entonces $ab = x_1x_2$. Ambos factores, x_1 y x_2 pertenecen a W, sin embargo $x_1 > x_2$. Por lo tanto aplicamos la tercera relación de (3.21): $x_1x_2 = [x_1, x_2] + P(x_1, x_2)x_2x_1$. Así, encontramos la representación deseada

$$x_1x_2 = [x_1, x_2] + P(x_1, x_2)x_2x_1 = w_{1,2} + q^{-1}w_{2,2}w_{1,1}.$$

El producto $ba = x_2x_1 = w_{2,2}w_{1,1}$ no necesita la aplicación del algoritmo, puesto que ya es monótono y $w_{2,2} < w_{1,1}$.

(ii) Consideremos un ejemplo con un ligero grado de dificultad mayor: $a = [x_1, x_2]x_1$, $b = x_2[x_1, x_2]$.

Tenemos que $ab = [x_1, x_2]x_1x_2[x_1, x_2]$. Este producto no es monótono: tiene la subpalabra x_1x_2 con $x_1 > x_2$.

En resultados anteriores se obtuvo que $x_1x_2 = [x_1, x_2] + q^{-1}x_2x_1$. Por lo tanto, $ab = [x_1, x_2][x_1, x_2][x_1, x_2] + [x_1, x_2]q^{-1}x_2x_1[x_1, x_2]$. El primer sumando es igual a $[x_1, x_2]^3 = w_{1,2}^3$ y es un producto monótono. El segundo sumando tiene dos palabras prohibidas: $[x_1, x_2]x_2$ y $x_1[x_1, x_2]$.

Por (3.21) tenemos que $[[x_1, x_2]x_2] = 0$, y $[x_1[x_1, x_2]] = 0$; esto es

$$\begin{aligned}[x_1, x_2]\, x_2 &= P(x_1x_2, x_2)x_2\, [x_1, x_2] \\ &= P(x_1, x_2)P(x_2, x_2)x_2\, [x_1, x_2] \\ &= qx_2\, [x_1, x_2]\,, \end{aligned} \tag{3.22}$$

y

$$\begin{aligned}x_1\, [x_1, x_2] &= P(x_1, x_1x_2)\, [x_1, x_2]\, x_1 \\ &= q\, [x_1, x_2]\, x_1. \end{aligned} \tag{3.23}$$

Puesto que $P(x_1x_2, x_2) = P(x_1, x_2)P(x_2, x_2) = q$, y $P(x_1, x_1x_2) = q$, tenemos la descomposición deseada

$$ab = [x_1, x_2]x_1x_2[x_1, x_2] = [x_1, x_2]^3 + qx_2[x_1, x_2]^2x_1.$$

(iii) Finalmente, presentamos un caso que requiere de más cálculos. Sea $N = 3$, $a = [x_1, x_2]\, x_1$ y $b = x_2^2\, [x_1, x_2]\, x_1$. Entonces $ab = [x_1, x_2]\, x_1x_2^2\, [x_1, x_2]\, x_1$. Puesto que $x_1 > x_2$, ab no es un producto monótono y procedemos a aplicar el algoritmo. Sabemos que $x_1x_2 = [x_1, x_2] + q^{-1}x_2x_1$ y entonces

$$\begin{aligned}ab &= [x_1, x_2]\, \left[[x_1, x_2] + q^{-1}x_2x_1\right] x_2\, [x_1, x_2]\, x_1 \\ &= [x_1, x_2]\, [x_1, x_2]\, x_2\, [x_1, x_2]\, x_1 + [x_1, x_2]\, q^{-1}x_2x_1x_2\, [x_1, x_2]\, x_1. \end{aligned}$$

De donde identificamos nuevamente dos relaciones prohibidas: $[x_1, x_2] < x_2$ y $x_1 > x_2$. Sustituyendo (3.22) en ab y desarrollando tenemos

$$\begin{aligned}ab =\ & [x_1, x_2]\, qx_2\, [x_1, x_2]\, [x_1, x_2]\, x_1 + q^{-1}qx_2\, [x_1, x_2]\, \left[[x_1, x_2] + q^{-1}x_2x_1\right] [x_1, x_2]\, x_1 \\ =\ & [x_1, x_2]\, qx_2\, [x_1, x_2]\, [x_1, x_2]\, x_1 + x_2\, [x_1, x_2]\, [x_1, x_2]\, [x_1, x_2]\, x_1 \\ & + x_2\, [x_1, x_2]\, q^{-1}x_2x_1\, [x_1, x_2]\, x_1. \end{aligned}$$

Nuevamente, se presentan las relaciones prohibidas $[x_1, x_2] < x_2$ en el primer sumando; y $x_1 < [x_1, x_2]$, en el tercero. Considerando (3.23) y sustituyendo este valor y el obtenido en (3.22) en nuestro último cálculo para el producto ab, obtenemos:

$$ab = q^2 x_2 [x_1, x_2]^3 x_1 + x_2 [x_1, x_2]^3 x_1 + q x_2^2 [x_1, x_2]^2 x_1^2.$$

Puesto que $N = 3$, $q^2 x_2 [x_1, x_2]^3 x_1 = 0$ y $x_2 [x_1, x_2]^3 x_1 = 0$. Se concluye entonces que

$$ab = q x_2^2 [x_1, x_2]^2 x_1^2.$$

Note que $\alpha = q$ si y sólo si $N = 3$.

4. COTAS PARA CÓDIGOS SOBRE ANILLOS DE FROBENIUS Y DE PESOS HOMOGÉNEOS

"Un experto en la resolución de problemas debe estar dotado de dos cualidades incompatibles: una imaginación inquieta y una paciente obstinación".
HOWARD W. EVES.

Entre las cotas clásicas en la teoría de códigos las más destacadas a mencionar son la cota de Hamming (conocida también en inglés como "sphere packing bound"), la cota de Singleton, la cota de Plotkin, la cota de Elias, la cota de programación lineal y la cota de Gilbert Varshamov.

Los códigos sobre anillos ganaron atención al inicio de la década de 1990 cuando se descubrió que ciertos códigos no lineales binarios de hecho tienen representación lineal en $\mathbb{Z}_4$ (véase Hammons˜Jr. y otros (1994)). Se ha observado que los códigos construidos sobre anillos son particulamente importantes cuando la estructura de la métrica en el alfabeto no está dada por el peso de Hamming, sino por variantes del mismo. En este sentido el llamado peso homogéneo, introducido en el artículo de Constantinescu y Heise (1997), resulta prominente. Por ejemplo, el peso de Lee en $\mathbb{Z}_4$ es homogéneo y códigos no lineales binarios y ternarios superiores han sido construidos a partir de códigos lineales sobre $\mathbb{Z}_8$ y $\mathbb{Z}_9$, en los que el anillo estaba dotado de pesos homogéneos (véase Duursma y otros [2001a, 2001b]).

Así como en el contexto de la teoría de códigos algebraica clásica sobre campos finitos, la pregunta natural en la teoría de códigos sobre anillos finitos es: ¿cuáles son los criterios para medir la calidad del código y qué códigos son óptimos? De ahí la necesidad de plantear analogías teóricas sobre anillos respecto a las citadas cotas.

A lo largo de este capítulo analizaremos de forma exhaustiva los resultados obtenidos en el artículo de Greferath y O'Sullivan (2004), mismo en que se establece la cota de Plotkin y en el que además se desarrollan versiones asintóticas de las cotas de Gilbert-Varshamov, de Elias y de Hamming, todas construidas sobre anillos de Frobenius con pesos homogéneos.

4.1. Fundamentos matemáticos

Este primer apartado tiene la finalidad de hacer una breve revisión sobre algunos conceptos matemáticos fundamentales necesarios para la cabal comprensión de las secciones subsecuentes.

4.1 Definición. Un conjunto G no vacío es un *grupo* si hay en G una operación binaria definida, llamada el producto y denotada por $\cdot$, tal que

(*i*) $a, b \in G$ implica que $a \cdot b \in G$ (cerradura).

(*ii*) $a, b, c \in G$ implica que $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (ley asociativa).

(*iii*) Existe un elemento $e \in G$ tal que $a \cdot e = e \cdot a = a$ para toda $a \in G$ (la existencia del elemento identidad en G).

(*iv*) Para toda $a \in G$ existe un elemento $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$ (la existencia de inversos en G).

4.2 Definición. Un subconjunto no vacío H de un grupo G se dice que es un *subgrupo* de G si, bajo el producto en G, H es un grupo.

4.3 Definición. Un mapeo ϕ del anillo R al anillo R' se dice un *homomorfismo* si

(*i*) $\phi(a + b) = \phi(a) + \phi(b)$ y

(*ii*) $\phi(ab) = \phi(a)\phi(b)$

para todas $a, b \in R$.

4.4 Definición. Si el homomorfismo de R a R' es uno a uno, se llama *isomorfismo*.

4.5 Definición. Si ϕ es un homomorfismo de R en R', entonces el *kernel* de ϕ, denotado como $ker(\phi)$, es el conjunto de todos los elementos $a \in R$ tal que $\phi(a) = 0$, los elementos cero en R'.

Por definición, el homomorfismo de R en R' es un isomorfismo si $ker(\phi) = \{0\}$.

4.6 Definición. Un grupo G se dice *abeliano* (o conmutativo) si para todo $a, b \in G$, $a \cdot b = b \cdot a$.

La noción de módulo puede extenderse considerando que los escalares no pertenecen únicamente a un campo, sino que simplemente son elementos de un anillo arbitrario.

4.7 Definición. Sea R un anillo cualquiera. Un conjunto no vacío M es un R-módulo (o un *módulo sobre* R) si M es un grupo abeliano bajo la operación + tal que para toda $r \in R$ y $m \in M$ existe un elemento $rm \in M$ que satisface:

(i) $r(a+b) = ra + rb$,

(ii) $r(sa) = (rs)a$, y

(iii) $(r+s)a = ra + sa$,

para toda $a, b \in M$ y $r, s \in R$.

Si R tiene el elemento unitario, 1, y si $1m = m$ para toda $m \in M$, entonces M se llama *R-módulo unital*. Notemos que en el caso en el que R es campo, un R-módulo unital no es más que un espacio vectorial sobre R.

Formalmente, cuando se considera un anillo no conmutativo la Definición 4.7 es la de un *R-módulo por la izquierda* y la definición de R-módulo por la derecha es análoga.

4.8 Definición. Si R es cualquier anillo, un subconjunto I_l de R se llama *idealo izquierdo* de R si

(i) I_l es un subgrupo de R bajo la adición.

(ii) $r \in R$, $a \in I_l$ implica $ra \in I_l$.

La definición de idealo derecho es análoga a la de idealo derecho.

4.9 Definición. Un idealo I que es idealo izquierdo e idealo derecho a la vez se llama *idealo*.

4.10 Ejemplo. Consideremos el anillo de los enteros $\mathbb{Z}$. Sea $A = \{\ldots, -4, -2, 0, 2, 4 \ldots\}$.

Claramente, A es un subgrupo de $\mathbb{Z}$ bajo la adición. Además, para todo entero r y a en A, se tiene que $ra \in A$. De hecho, cada elemento en A tiene la forma $2tr$. Por lo tanto, y por cumplirse la conmutatividad en los enteros, A es idealo de $\mathbb{Z}$.

4.2. Pesos Homogéneos y Anillos de Frobenius

El concepto de peso homogéneo se debe a Constantinescu y Heise (1997) quienes abordaron el caso especial en que $R = \mathbb{Z}_m$, el anillo de los enteros módulo m.

4.11 Definición. Una función real w en el anillo finito R se llama *peso homogéneo (izquierdo)* si $w(0) = 0$ y si satisface lo siguiente:

(i) Para toda $x, y \in R$ $Rx = Ry$ implica $w(x) = w(y)$.

(ii) Existe un número real γ tal que

$$\sum_{y \in Rx} w(y) = \gamma \, |Rx| \qquad \text{para toda } x \in R \setminus \{0\}.$$

Los pesos homogéneos derechos se definen de forma análoga, y de aquí en adelante nos referiremos a los pesos homogéneos izquierdos simplemente como pesos homogéneos. El número γ es el valor promedio de w en R, y la condición (ii) dice que el valor promedio de w es constante en cada idealo principal distinto de cero en R.

4.12 Proposición. Sea R un anillo de Frobenius finito con caracter generador χ. Entonces todo peso homogéneo en R es de la forma

$$w : R \to \mathbb{R}, \qquad x \mapsto \gamma \left[1 - \frac{1}{|R^\times|} \sum_{u \in R^\times} \chi(xu) \right].$$

Demostración. w es un mapeo del anillo de Frobenius R a los reales y $\mapsto$ enfatiza el efecto de w sobre el dominio R. Por la unicidad del peso homogéneo en un anillo fijo R, solamente es necesario verificar que la w dada en la proposición satisfaga los criterios de la Definición 4.11. Como $-1 \in R^\times$ se tiene que w es una función real y, ciertamente, $w(0) = 0$. Si $Rx = Ry$ entonces $y = vx$ para algún $v \in R^\times$. Por Wood (1999) sabemos que todo caracter generador izquierdo es también un caracter generador derecho. Por esta razón, existe $v' \in R^\times$ tal que $\chi(vx) = \chi(xv')$ y tenemos que

$$\sum_{u \in R^\times} \chi(vxu) = \sum_{u \in R^\times} \chi(xu),$$

lo que finalmente muestra que $w(x) = w(y)$. Para probar que la segunda parte de la definición se cumple, se observa que para toda $x \neq 0$ la expresión $\chi(Rxu)$ no es trivial puesto que el kernel de un caracter generador no contiene idealos distintos de cero de R. Por lo tanto

$$\sum_{y \in Rx} \frac{1}{|R^\times|} \sum_{u \in R^\times} \chi(yu) = \frac{1}{|R^\times|} \sum_{u \in R^\times} \sum_{y \in Rx} \chi(yu) = \frac{1}{|R^\times|} \sum_{u \in R^\times} \sum_{y \in Rxu} \chi(y) = 0,$$

lo cual completa la prueba. □

4.13 Proposición. Sea R un anillo de Frobenius finito, sea $w : R \to \mathbb{R}$ un peso homogéneo con valor promedio $\gamma \geq 0$, y sea P la distribución de probabilidad en R. Entonces, la siguiente desigualdad se cumple:

$$\sum_{x,y \in R} w(x-y)P(x)P(y) \leq \gamma.$$

Demostración. Escribiremos la expresión anterior con el peso homogéneo en términos de χ, tenemos

$$\begin{aligned}
\sum_{x,y \in R} w(x-y)P(x)P(y) &= \gamma \sum_{x,y \in R} \left[1 - \frac{1}{|R^\times|} \sum_{u \in R^\times} \chi\left[(x-y)u\right] \right] P(x)P(y) \\
&= \gamma \left[\sum_{x,y \in R} P(x)P(y) - \frac{1}{|R^\times|} \sum_{u \in R^\times} \sum_{x,y \in R} \chi\left[xu - yu\right] P(x)P(y) \right] \\
&= \gamma \left[\sum_{x,y \in R} P(x)P(y) - \frac{1}{|R^\times|} \sum_{u \in R^\times} \sum_{x,y \in R} \chi(xu)\overline{\chi(yu)}P(x)P(y) \right] \\
&= \gamma \left[1 - \frac{1}{|R^\times|} \sum_{u \in R^\times} \sum_{x,y \in R} \chi(xu)\overline{\chi(yu)}P(x)P(y) \right] \\
&= \gamma \left[1 - \frac{1}{|R^\times|} \sum_{u \in R^\times} \left| \sum_{x \in R} \chi(xu)P(x) \right|^2 \right] \leq \gamma
\end{aligned}$$

lo cual prueba la proposición. □

De la proposición anterior, y considerando una distribución uniforme para los subconjuntos de R, obtenemos el siguiente corolario:

4.14 Corolario. Sea R un anillo de Frobenius finito y sea $w : R \to \mathbb{R}$ un peso homogéneo con valor promedio $\gamma \geq 0$. Para cada subconjunto I de R se cumple que:

$$\sum_{x,y \in I} w(x-y) \leq \gamma \left| I \right|^2.$$

Demostración.

$$\sum_{x,y\in I} w(x-y)P(x)P(y) \leq \gamma$$
$$\sum_{x,y\in I} w(x-y)\frac{1}{|I|}\frac{1}{|I|} \leq \gamma$$
$$\sum_{x,y\in I} w(x-y) \leq \gamma\,|I|^2\,.$$

□

La siguiente proposición es una versión refinada de la Proposición 4.13.

4.15 Proposición. Sean R un anillo de Frobenius finito; $w : R \to \mathbb{R}$, un peso homogéneo de valor promedio $\gamma \geq 0$; y P, la distribución de probabilidad en R. Entonces, la siguiente desigualdad se satisface:

$$\sum_{x,y\in R} w(x-y)P(x)P(y) \leq 2\sum_{x\in R} w(x)P(x) - \frac{1}{\gamma}\left[\sum_{x\in R} w(x)P(x)\right]^2.$$

Demostración. De la Proposición 4.13 sabemos que

$$\sum_{x,y\in R} w(x-y)P(x)P(y) = \gamma\left[1 - \frac{1}{|R^\times|}\sum_{u\in R^\times}\left|\sum_{x\in R}\chi(xu)P(x)\right|^2\right].$$

Conviene recordar el Lema 2.21:

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

donde $\{a_1,\ldots,a_n\}$ y $\{b_1,\ldots,b_n\}$ son dos conjuntos de números reales.

Haciendo la suma sobre toda $u \in R^\times$, $a_u = \frac{1}{|R^\times|}$ y $b_u = \sum_{x\in R}\chi(xu)P(x)$ en la desigualdad de Cauchy-Schwarz, tenemos

$$\left(\sum_{u\in R^\times}\frac{1}{|R^\times|}\sum_{x\in R}\chi(xu)P(x)\right)^2 \leq \left(\sum_{u\in R^\times}\frac{1}{|R^\times|^2}\right)\left(\sum_{u\in R^\times}\left(\sum_{x\in R}\chi(xu)P(x)\right)^2\right).$$

Desarrollando la desigualdad

$$\begin{aligned}\frac{1}{|R^\times|^2}\left(\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right)^2 &\leq \frac{1}{|R^\times|^2}\left(\sum_{u\in R^\times}1\right)\left(\sum_{u\in R^\times}\left(\sum_{x\in R}\chi(xu)P(x)\right)^2\right)\\ &= \frac{1}{|R^\times|^2}\left|R^\times\right|\left(\sum_{u\in R^\times}\left(\sum_{x\in R}\chi(xu)P(x)\right)^2\right)\\ &= \frac{1}{|R^\times|}\sum_{u\in R^\times}\left(\sum_{x\in R}\chi(xu)P(x)\right)^2\end{aligned}$$

y puesto que

$$\begin{aligned}\left(\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right)^2 &= \left|\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right|^2\\ \left(\sum_{x\in R}\chi(xu)P(x)\right)^2 &= \left|\sum_{x\in R}\chi(xu)P(x)\right|^2\end{aligned}$$

obtenemos la desigualdad

$$\frac{1}{|R^\times|^2}\left|\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right|^2 \leq \frac{1}{|R^\times|}\sum_{u\in R^\times}\left|\sum_{x\in R}\chi(xu)P(x)\right|^2. \tag{4.1}$$

Ahora, desarrollamos el primer miembro de la desigualdad considerando además que,

por la ecuación (4.12), $1 - \frac{w(x)}{\gamma} = \frac{1}{|R^\times|}\sum_{u\in R^\times}\chi(xu)$:

$$\begin{aligned}
\frac{1}{|R^\times|^2}\left|\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right|^2 &= \left|\frac{1}{|R^\times|}\sum_{u\in R^\times}\sum_{x\in R}\chi(xu)P(x)\right|^2 \\
&= \left|\sum_{x\in R}P(x)\frac{1}{|R^\times|}\sum_{u\in R^\times}\chi(xu)\right|^2 \\
&= \left|\sum_{x\in R}P(x)\left(1-\frac{w(x)}{\gamma}\right)\right|^2 \\
&= \left|\sum_{x\in R}P(x) - \frac{1}{\gamma}\sum_{x\in R}w(x)P(x)\right|^2 \\
&= \left[1-\frac{1}{\gamma}\sum_{x\in R}w(x)P(x)\right]^2 \\
&= 1 - \frac{2}{\gamma}\sum_{x\in R}w(x)P(x) + \frac{1}{\gamma^2}\left[\sum_{x\in R}w(x)P(x)\right]^2.
\end{aligned}$$

De manera que la desigualdad (4.1) puede escribirse como

$$\begin{aligned}
\frac{1}{|R^\times|}\sum_{u\in R^\times}\left|\sum_{x\in R}\chi(xu)P(x)\right|^2 &\geq 1 - \frac{2}{\gamma}\sum_{x\in R}w(x)P(x) + \frac{1}{\gamma^2}\left[\sum_{x\in R}w(x)P(x)\right]^2 \\
\gamma\left[\frac{1}{|R^\times|}\sum_{u\in R^\times}\left|\sum_{x\in R}\chi(xu)P(x)\right|^2 - 1\right] &\geq -2\sum_{x\in R}w(x)P(x) + \frac{1}{\gamma}\left[\sum_{x\in R}w(x)P(x)\right]^2 \\
\gamma\left[1-\frac{1}{|R^\times|}\sum_{u\in R^\times}\left|\sum_{x\in R}\chi(xu)P(x)\right|^2\right] &\leq 2\sum_{x\in R}w(x)P(x) - \frac{1}{\gamma}\left[\sum_{x\in R}w(x)P(x)\right]^2
\end{aligned}$$

lo que completa la prueba. □

4.3. La Cota de Plotkin

Sea C un código no necesariamente lineal sobre el anillo de Frobenius R, mismo que cuenta con el peso homogéneo w y con valor promedio $\gamma \geq 0$. Los parámetros del código están dados por (n, M, d), donde n es la longitud del código; $M = |C|$, su cardinalidad; y su distancia mínima w está dada por d. Recordemos que deseamos encontrar, en este

caso, una cota superior para

$$A_w(n,d) = \max\{M : \text{existe un código } (n,M,d) \text{ sobre } R\}.$$

La siguiente proposición es una generalización de la cota de Plotkin para códigos sobre campos finitos, caso clásico de estudio de la teoría de códigos que analizamos en el Capítulo 2: *Cotas en Teoría de Códigos*.

4.16 Proposición. Para todo código $C-(n,M,d)$ se satisface que

$$M(M-1)d \leq \sum_{\mathbf{x},\mathbf{y}\in C} w(\mathbf{x}-\mathbf{y}) \leq \gamma n M^2.$$

Demostración. La primera parte de la desigualdad se sigue del hecho de que $w(\mathbf{x}-\mathbf{y}) \geq d$ para $x \neq y$.[1] La segunda parte de la desigualdad se obtiene a través de un proceso de marginalización que a continuación se explica. Para cualquier distribución de probabilidad P sobre R^n consideremos la siguiente correspondencia:

$$\pi_i : R^n \longrightarrow R \qquad x \mapsto x_i. \tag{4.2}$$

La proyección a la i-ésima coordenada induce una distribución P_i en R. Dicho de otra forma, en lugar de considerar las distribuciones de probabilidad por cada palabra código, se considerarán las distribuciones de probabilidad por cada "letra" de cada palabra código. Por esta razón, aplicamos la Proposición 4.13 para obtener

$$\begin{aligned}\sum_{\mathbf{x},\mathbf{y}\in R^n} w(\mathbf{x}-\mathbf{y})P(\mathbf{x})P(\mathbf{y}) &= \sum_{i=1}^{n}\sum_{\mathbf{x},\mathbf{y}\in R^n} w(x_i-y_i)P(\mathbf{x})P(\mathbf{y}) \\ &= \sum_{i=1}^{n}\sum_{r,s\in R} w(r-s)P_i(r)P_i(s) \leq \sum_{i=1}^{n}\gamma = \gamma n.\end{aligned}$$

Asumiendo una distribución de probabilidad uniforme sobre C, es decir, $P(\mathbf{x}) = \frac{1}{M}$ si $\mathbf{x} \in C$ y $P(\mathbf{x}) = 0$ en otro caso, se tiene el segundo miembro de la desigualdad, completando así la prueba. □

4.17 Teorema (Cota de Plotkin)**.** Para toda n, d con $\gamma n \leq d$ se cumple que

$$A_w(n,d) \leq \frac{d}{d-\gamma n}.$$

Demostración. La prueba se sigue directamente, por transitividad, de la Proposición 4.16. □

[1] Véase la deducción de la desigualdad (2.3), pues $d(\mathbf{x},\mathbf{y}) = w(\mathbf{x}-\mathbf{y})$.

4.4. La Cota de Elias

De acuerdo con Greferath y O'Sullivan (2004) esta cota está relacionada estrechamente a la cota de Plotkin. Puede ser considerada como una versión refinada de la cota de Plotkin basada, en el caso de códigos sobre anillos de Frobenius, en la Proposición 4.15, como veremos en breve.

4.18 Proposición. Para las distribuciones de probabilidad P en R^n, se tiene que

$$\sum_{\mathbf{x},\mathbf{y}\in R^n} w(\mathbf{x}-\mathbf{y})P(\mathbf{x})P(\mathbf{y}) \leq 2\sum_{\mathbf{x}\in R^n} w(\mathbf{x})P(\mathbf{x}) - \frac{1}{\gamma n}\left[\sum_{\mathbf{x}\in R^n} w(\mathbf{x})P(\mathbf{x})\right]^2.$$

Demostración. Combinando la prueba de la Proposición 4.16 y la Proposición 4.15 se sigue que

$$\begin{aligned}\sum_{\mathbf{x},\mathbf{y}\in R^n} w(\mathbf{x}-\mathbf{y})P(\mathbf{x})P(\mathbf{y}) &= \sum_{i=1}^{n}\sum_{r,s\in R} w(r-s)P_i(r)P_i(s) \\ &\leq \sum_{i=1}^{n}\left[2\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma}\left[\sum_{r\in R} w(r)P_i(r)\right]^2\right].\end{aligned}$$

Retomemos nuevamente la desigualdad de Cauchy-Schwarz. Haciendo correr la suma sobre $1 \leq i \leq n$, $a_i = 1$ y $b_i = \sum_{r\in R} w(r)P_i(r)$ y al sustituir estos valores en la ecuación de Cauchy-Schwarz:

$$\begin{aligned}\left(\sum_{i=1}^{n} 1\sum_{r\in R} w(r)P_i(r)\right)^2 &\leq \left(\sum_{i=1}^{n} 1^2\right)\left(\sum_{i=1}^{n}\left(\sum_{r\in R} w(r)P_i(r)\right)^2\right) \\ &\leq n\sum_{i=1}^{n}\left[\sum_{r\in R} w(r)P_i(r)\right]^2\end{aligned}$$

esto es

$$\sum_{i=1}^{n}\left[\sum_{r\in R} w(r)P_i(r)\right]^2 \geq \frac{1}{n}\left[\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r)\right]^2$$

de donde se tiene que

$$-\frac{1}{\gamma}\sum_{i=1}^{n}\left[\sum_{r\in R} w(r)P_i(r)\right]^2 \leq -\frac{1}{\gamma n}\left[\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r)\right]^2.$$

Sumando a cada miembro de esta última desigualdad $2\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r)$:

$$2\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma}\sum_{i=1}^{n}\left[\sum_{r\in R} w(r)P_i(r)\right]^2 \leq 2\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma n}\left[\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r)\right]^2,$$

y además

$$\begin{aligned}\sum_{\mathbf{x},\mathbf{y}\in R^n} w(\mathbf{x}-\mathbf{y})P(\mathbf{x})P(\mathbf{y}) &\leq \sum_{i=1}^{n}\left[2\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma}\left[\sum_{r\in R} w(r)P_i(r)\right]^2\right] \\ &= 2\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma}\sum_{i=1}^{n}\left[\sum_{r\in R} w(r)P_i(r)\right]^2,\end{aligned}$$

por lo que, por transitividad, se obtiene

$$\begin{aligned}\sum_{\mathbf{x},\mathbf{y}\in R^n} w(\mathbf{x}-\mathbf{y})P(\mathbf{x})P(\mathbf{y}) &\leq 2\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r) - \frac{1}{\gamma n}\left[\sum_{i=1}^{n}\sum_{r\in R} w(r)P_i(r)\right]^2 \\ &= 2\sum_{\mathbf{x}\in R^n} w(\mathbf{x})P(\mathbf{x}) - \frac{1}{\gamma n}\left[\sum_{\mathbf{x}\in R^n} w(\mathbf{x})P(\mathbf{x})\right]^2.\end{aligned}$$

□

4.19 Corolario. Sea C un código con parámetros (n, M, d) y $T := \sum_{\mathbf{c}\in C} w(\mathbf{c})$ el *peso total* de C. Entonces, se cumple que

$$\sum_{\mathbf{x},\mathbf{y}\in C} w(\mathbf{x}-\mathbf{y}) \leq 2MT - \frac{1}{\gamma n}T^2.$$

Demostración. $|C| = M$ y aplicando una distribución uniforme en C, y cero fuera de C, obtenemos directamente de la proposición anterior:

$$\frac{1}{M^2}\sum_{\mathbf{x},\mathbf{y}\in C} w(\mathbf{x}-\mathbf{y}) \leq 2\frac{T}{M} - \frac{1}{\gamma n}\frac{T^2}{M^2}.$$

La prueba queda completada al multiplicar esta última expresión por M^2. □

4.20 Corolario. Sea C un código con parámetros (n, M, d) y sea $W := max_{\mathbf{c}\in C} w(\mathbf{c})$ el *peso máximo* de C. Si $W \leq \gamma n$ y $d\gamma n - 2W\gamma n + W^2 > 0$, entonces

$$M \leq \frac{d\gamma n}{d\gamma n - 2W\gamma n + W^2}.$$

Demostración. Consideremos la función

$$f(z) = 2Mz - \frac{z^2}{\gamma n}.$$

Sabemos que una función $f(z)$ es creciente si y sólo si $f'(z) \geq 0$. Puesto que

$$f'(z) = 2M - \frac{2z}{\gamma n} \geq 0$$

implica $M\gamma n \geq z$, la función $f(z)$ es creciente en el intervalo $(-\infty, M\gamma n]$. La conveniencia de usar esta función radica en el hecho de que $f(M\gamma n) = M^2\gamma n$, la cota superior de la Proposición 4.16.

Por otra parte, es claro que el peso total T de C satisface $T \leq MW$. Puesto que f es creciente, si $MW \leq M\gamma n$, entonces $f(T) \leq f(MW)$, y en combinación con la Proposición 4.16:

$$M(M-1)d \leq \sum_{\mathbf{x},\mathbf{y}\in C} w(\mathbf{x}-\mathbf{y}) \leq 2M^2W - \frac{M^2W^2}{\gamma n}.$$

De manera que, nuevamente por transitividad, $M(M-1)d \leq 2M^2W - \frac{M^2W^2}{\gamma n}$ y despejando a M se obtiene el resultado deseado. □

Cabe destacar que en la desigualdad del Corolario 4.20, W puede reemplazarse por $W := min_{\mathbf{z}\in R^n} max_{\mathbf{c}\in C} w(\mathbf{c}-\mathbf{z})$. Esto es, W puede reemplazarse por el peso máximo de cualquier traducción de C.

4.21 Proposición. Sea $C-(n, M, d)$ un código sobre R . Entonces, para toda $t \geq 0$ existe un elemento $\mathbf{z} \in R^n$ tal que la esfera $S_{R_w}(\mathbf{z}, t)$, de radio t y centro en $\mathbf{z}$, satisface

$$|C \cap S_{R_w}(\mathbf{z}, t)| \geq \frac{M \cdot V_{R_w}(n, t)}{|R|^n}.$$

donde $V_{R_w}(n, t)$ denota el volumen de la esfera $S_{R_w}(\mathbf{z}, t)$ (véase la Definición 2.12).

Demostración. Contemos los pares $(\mathbf{c}, \mathbf{z})$ tales que $\mathbf{c} \in C$ y $\mathbf{z} \in R^n$, considerando que $\mathbf{c}$ está en la esfera $S_{R_w}(\mathbf{z}, t)$. Por una parte, por cada $\mathbf{c} \in C$ existen $V_{R_w}(n, t)$ elementos $\mathbf{z} \in R^n$ tales que $\mathbf{c} \in S_{R_w}(\mathbf{z}, t)$, por lo que tenemos un total de $M \cdot V_{R_w}(n, t)$ pares $(\mathbf{c}, \mathbf{z})$. Por otra parte, también es cierto que por cada $\mathbf{z} \in R^n$, el conjunto de palabras código tales que $\mathbf{c} \in S_{R_w}(\mathbf{z}, t)$ está dado por la intersección del código y la esfera: $C \cap S_{R_w}(\mathbf{z}, t)$, por lo que se tiene un total de $\sum_{z\in R^n} |C \cap S_{R_w}(\mathbf{z}, t)|$ pares $(\mathbf{c}, \mathbf{z})$. Puesto que ambas

expresiones son válidas para contar los pares $(\mathbf{c}, \mathbf{z})$, y dividiendo ambas expresiones por $|R|^n$, se concluye que

$$\sum_{z \in R^n} \frac{|C \cap S_{R_w}(\mathbf{z}, t)|}{|R|^n} = \frac{M \cdot V_{R_w}(n, t)}{|R|^n}.$$

Es claro que se tiene entonces un promedio, por lo que siempre existe un elemento $\mathbf{z} \in R^n$ tal que

$$|C \cap S_{R_w}(\mathbf{z}, t)| \geq \frac{M \cdot V_{R_w}(n, t)}{|R|^n}.$$

□

Con todos estos conceptos podemos ahora construir la cota de Elias para códigos sobre anillos de Frobenius finitos equipados con pesos homogéneos.

4.22 Teorema (Cota de Elias)**.** Para toda n, d, t con $t \leq \gamma n$ y $t^2 - 2t\gamma n + d\gamma n > 0$ se cumple que

$$A_w(n, d) \leq \frac{\gamma n d}{t^2 - 2t\gamma n + d\gamma n} \frac{|R|^n}{V_{R_w}(n, t)}.$$

Demostración. Sea C un código con parámetros (n, M, d) definido sobre el anillo de Frobenius R equipado con pesos homogéneos w y, para toda $t > 0$ elijamos un centro $\mathbf{z}_t$ para una esfera como la que explicamos en la proposición anterior. Sean $C_t := C \cap S_{R_w}(\mathbf{z}_t, t)$ y (n, M_t, d_t) los parámetros de C_t. De manera que, por ser d minimal, $d \leq d_t$ y el radio de la esfera t satisface $t \geq min_{\mathbf{z} \in R^n} max_{\mathbf{c} \in C_t} w(\mathbf{c} - \mathbf{z})$ y por lo tanto, de la última ecuación en la prueba del Corolario 4.20, obtenemos la desigualdad

$$M_t(M_t - 1)d \leq M_t(M_t - 1)d_t \leq \sum_{\mathbf{x}, \mathbf{y} \in C_t} w(\mathbf{x} - \mathbf{y}) \leq M_t^2 \left[2t - \frac{t^2}{\gamma n}\right].$$

Por transitividad, y despejando M_t tenemos que

$$M_t \leq \frac{d\gamma n}{d\gamma n - 2t\gamma n + t^2} \tag{4.3}$$

puesto que $t \leq \gamma n$ y el denominador es positivo. Por la Proposición 4.21, y considerando que $|C_t| = |C \cap S_{R_w}(\mathbf{z}_t, t)| = M_t$ obtenemos

$$M \leq \frac{M_t |R|^n}{V_{R_w}(n, t)}, \tag{4.4}$$

y, finalmente, combinando (4.3) y (4.4):

$$M \leq \frac{M_t |R|^n}{V_{R_w}(n, t)} \leq \frac{d\gamma n}{d\gamma n - 2t\gamma n + t^2} \frac{|R|^n}{V_{R_w}(n, t)}.$$

lo cual completa la prueba. □

4.23 Ejemplo. En el anillo de los enteros $R = \mathbb{Z}_8$, el conjunto de unidades es $R^\times = \{1, 3, 5, 7\}$ e $I = \{0, 2, 4, 6\}$ es idealo del anillo, como puede observarse en la tabla del producto de $\mathbb{Z}_8$ (Tabla 4.1). Definamos una función de pesos homogéneos de la siguiente forma:

$$w_{hom} : \mathbb{Z}_8 \to, r \mapsto \begin{cases} 0 & \text{si } r = 0; \\ 2 & \text{si } r = 4; \\ 1 & \text{en otro caso.} \end{cases}$$

Evaluaremos la cota de Elias para un código con parámetros $n = 18$ y $d = 8$.

Por el Teorema 4.22, t debe satisfacer $t \leq 18$ y $t^2 - 36t + 144 > 0$. Sabemos que la solución de $t^2 + pt + q = 0$ está dada por la fórmula

$$t_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

Haciendo $p = -36$ y $q = 144$ se tiene que $t_1 \approx 4.58$ y $t_2 \approx 31.42$. Puesto que en este caso se trata de una desigualdad, y además se consideran valores enteros para t, se observa que $t_1 \leq \lfloor 4.58 \rfloor = 4$ y $t_2 \geq \lceil 31.42 \rceil = 32$. Debido a la restricción $t \leq 18$, los valores de t_2 quedan descartados y tenemos entonces sólo cinco posibles valores de t: 0,1 2 3 y 4. Por supuesto, el valor deseado de t es aquel que minimiza

$$\frac{\gamma n d}{t^2 - 2t\gamma n + d\gamma n} \frac{|R|^n}{V_{R_w}(n, t)}.$$

Recordemos que, conforme a lo expuesto durante el Capítulo 2, para enteros positivos dados $q > 1$, n y $t \geq 0$, $V_w(n, t)$ el volumen de una esfera con radio t en R^n se define como

$$V_{R_w}(n, t) = \begin{cases} \binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \ldots + \binom{n}{t}(q-1)^t & \text{si } 0 \leq t \leq n; \\ q^n & \text{si } n \leq t. \end{cases}$$

Con el fin de simplificar los cálculos, haremos uso de logaritmos en base 8 y del cambio de base dado por

$$\log_8(a) = \frac{\log_{10}(a)}{\log_{10}(8)}.$$

La cota de Elias para cada valor posible de t es entonces:

$$t = 0,\ \log_8(A_w(18, 8)) \leq 18;$$
$$t = 1,\ \log_8(A_w(18, 8)) \leq 15.80;$$

×	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	0	2	4	6
3	0	3	6	1	4	7	2	5
4	0	4	0	4	0	4	0	4
5	0	5	2	7	4	1	6	3
6	0	6	4	2	0	6	4	2
7	0	7	6	5	4	3	2	1

Tab. 4.1: Tabla del producto en $\mathbb{Z}_8$

$$t = 2,\ \log_8(A_w(18,8)) \leq 14.01;$$
$$t = 3,\ \log_8(A_w(18,8)) \leq 12.51;$$

y

$$t = 4,\ \log_8(A_w(18,8)) \leq 11.44 = \log_8\left(\frac{144}{16 \cdot 7634572}\right) + 18,$$

siendo $t = 4$ el valor que genera la mejor cota del ejemplo.

4.5. Cotas asintóticas

Sea w un peso homogéneo de valor promedio $\gamma \geq 0$ en el anillo finito de Frobenius R. Sea $\mathbb{R}_+$ el conjunto de reales no negativos. Definimos el anillo monoide como

$$\mathbb{Z}[\mathbb{R}_+] := \left\{ \sum_{i=1}^{n} f_i z^{r_i} | f_i \in \mathbb{Z}, r_i \in \mathbb{R}_+ \text{ y } n \in \mathbb{N} \right\},$$

donde z es un *indeterminado*. La función generadora w es el elemento $f_w(z) = \sum_{r \in R} z^{w(r)}$ de $\mathbb{Z}[\mathbb{R}_+]$. Es claro que el peso promedio en R está dado por

$$\gamma := \frac{1}{|R|} \sum_{r \in R} w(r)$$

y además $\frac{d}{dz} f_w(z) = \frac{d}{dz}\left[\sum_{r \in R} z^{w(r)}\right] = \sum_{r \in R} w(r) z^{w(r)-1}$, de donde se tiene que

$$\frac{d}{dz} f_w(1) = \sum_{r \in R} w(r)$$

y por tanto

$$\gamma := \frac{1}{|R|}\sum_{r\in R} w(r) = \frac{1}{|R|}\frac{d}{dz} f_w(1).$$

Ahora, estudiaremos el comportamiento asintótico de la esfera de radio δn para $n \in \mathbb{N}$ y $\delta \in [0,1]$.

4.24 Teorema. Para toda $\delta \in [0,\gamma]$, se satisface que

$$\lim_{n\to\infty} \frac{1}{n} \log_{|R|} V_w(n,\delta n) = min_{z\in(0,1]} \log_{|R|} \frac{f_w(z)}{z^\delta}.$$

Demostración. El artículo de Loeliger (1994) precisa la función

$$\pounds(\lambda) := \sum_{r\in R} \exp(-\lambda w(r))$$

y se prueba que

$$V_{R_w}(n,\delta n) \leq \left[\exp(\lambda\delta)\pounds(\lambda)\right]^n,$$

para toda $\lambda \in [0,\infty)$ y $n \in \mathbb{N}$. Sustituyendo $z = \exp(-\lambda)$ este rango se transforma a $z \in (0,1]$ y $\pounds(\lambda)$ puede escribirse como $f_w(z)$:

$$z^{-\delta} = exp(\lambda\delta),$$

$$\sum_{r\in R} \exp(-\lambda w(r)) = \sum_{r\in R} z^{w(r)}.$$

Por lo tanto, obtenemos

$$V_{R_w}(n,\delta n) \leq \left[\frac{f_w(z)}{z^\delta}\right]^n$$

para toda $z \in (0,1]$ y $n \in \mathbb{N}$, y por medio de logaritmos

$$\frac{1}{n}\log_{|R|} V_{R_w}(n,\delta n) \leq \min_{z\in(0,1]} \log_{|R|} \frac{f_w(z)}{z^\delta}$$

para toda $n \in \mathbb{N}$. El comentario que sigue a la ecuación (4) del artículo de Loeliger (1994) indica que esta desigualdad tiene un comportamiento asintótico. □

4.25 Definición. Para un anillo de Frobenius R, un peso w definido sobre R con función generadora f_w y un número real no negativo δ, abreviamos

$$h_w(\delta) := min_{z\in(0,1]} \log_{|R|} \frac{f_w(z)}{z^\delta}.$$

Sabemos que el mínimo de una función se encuentra cuando su primera derivada es cero. Hacemos

$$\frac{d}{dz}\log_{|R|}\frac{f_w(z)}{z^\delta}=0,$$

y, por tratarse de una función creciente puesto que $|R|>1$, basta con calcular la derivada únicamente sobre el argumento de la función logarítmica

$$\frac{d}{dz}\frac{f_w(z)}{z^\delta}=\frac{z^\delta\frac{d}{dz}f_w(z)-f_w(z)\cdot\delta z^{\delta-1}}{z^{2\delta}}=0.$$

Esto es, el punto z que resuelve la expresión $z\frac{d}{dz}f_w(z)=\delta f_w(z)$ es el mínimo.

4.26 Ejemplo. Consideremos el campo finito F_q con $q=3$ equipado con el peso de Hamming w_H. Sabemos que el peso de Hamming para cada elemento en el campo F_3 está dado como sigue: $w_H(0)=0$ y $w_H(1)=w_H(2)=1$. Primero, calcularemos $f_w(z)$ y γ:

$$\begin{aligned}f_w(z)&=\sum_{r\in F_3}z^{w_H(r)}\\&=z^{w_H(0)}+z^{w_H(1)}+z^{w_H(2)}\\&=z^0+z^1+z^1\\&=1+2z.\end{aligned}$$

$$\gamma=\frac{\sum_{r\in F_3}w_H(r)}{|F_3|}=\frac{0+1+1}{3}=\frac{2}{3}.$$

El mínimo de la función $f_{w_H}(z)$ está dado por el punto z tal que

$$\begin{aligned}z\frac{d}{dz}f_{w_H}(z)&=\delta f_{w_H}(z)\\z\frac{d}{dz}(1+2z)&=\delta(1+2z)\\2z&=\delta+2\delta z\end{aligned}$$

de donde se tiene que

$$z = \frac{\delta}{2 - 2\delta}.$$

Ahora, podemos calcular $h_{w_H}(\delta)$:

$$\begin{aligned} h_{w_H}(\delta) &= \min_{z\in(0,1]} \log_3 \frac{f_{w_H}(z)}{z^\delta} \\ &= \min_{z\in(0,1]} \left(\log_3 f_{w_H}(z) - \log_3 z^\delta\right) \\ &= \min_{z\in(0,1]} \left(\log_3(1+2z) - \delta \log_3 z\right) \\ &= \log_3\left(1 + \frac{2\delta}{2-2\delta}\right) - \delta \log_3\left(\frac{\delta}{2-2\delta}\right) \\ &= \delta \log_3 \frac{1}{\delta} + (1-\delta)\log_3\left(\frac{1}{1-\delta}\right) + \delta \log_3(2). \end{aligned}$$

4.27 Ejemplo. Generalicemos el Ejemplo 4.26. Sea F_q un campo finito con q elementos equipado con el peso de Hamming w_H. Puesto que cada elemento en F_q tiene un peso de Hamming igual a 1, excepto el 0, se tiene que $f_{w_H}(z) = 1 + (q-1)z$ y $\gamma = (q-1)/q$. De forma análoga al ejemplo anterior, resolvemos $z(d/dz)f_{w_H}(z) = \delta f_{w_H}(z)$, y obtenemos $z := \delta/(q-1)(1-\delta)$. Asimismo obtenemos para toda $0 \le \delta \le \gamma$ la ecuación:

$$\begin{aligned} h_{w_H}(\delta) &= \min_{z\in(0,1]} \log_q \frac{f_{w_H}(z)}{z^\delta} \\ &= \min_{z\in(0,1]} \left(\log_q f_{w_H}(z) - \log_q z^\delta\right) \\ &= \min_{z\in(0,1]} \left(\log_q(1+(q-1)z) - \delta \log_q z\right) \\ &= \log_q\left(1 + (q-1)\frac{\delta}{(q-1)(1-\delta)}\right) - \delta \log_q\left(\frac{\delta}{(q-1)(1-\delta)}\right) \\ &= \log_q\left(\frac{1}{1-\delta}\right) - \delta \log_q \delta + \delta \log_q(q-1) + \delta \log_q(q-1) \\ &= \delta \log_q \frac{1}{\delta} + (1-\delta)\log_q \frac{1}{1-\delta} + \delta \log_q(q-1). \end{aligned}$$

En lo subsecuente, buscaremos cotas asintóticas para la cantidad

$$\alpha_w(\delta) := \limsup_{n\to\infty} \frac{1}{n} \log_{|R|} A_w(n, \delta n),$$

que es el límite superior de la mayor tasa posible de un código de longitud n y distancia mínima relativa δn.

El siguiente teorema hace uso de las cotas de Gilbert-Varshamov y de Hamming, en sus versiones asintóticas.

4.28 Teorema. *Sea w un peso homogéneo de valor promedio γ en el anillo de Frobenius finito R. Para toda $\delta \in [0, \gamma]$*

$$1 - h_w(\delta) \leq \alpha_w(\delta) \leq 1 - h_w\left(\frac{\delta}{2}\right).$$

Demostración. Para un código $C - (n, M, d)$ la cota de la esfera establece que

$$M \cdot V_w(n, (d-1)/2) \leq |R|^n.$$

Sea $d = \delta n$. Asintóticamente, tenemos que

$$\limsup_{n\to\infty} \frac{1}{n} \log_{|R|} A_w(n, \delta n) \leq \limsup_{n\to\infty} \frac{1}{n} \log_{|R|} \frac{|R|^n}{V_w(n, (\delta n - 1)/2)}.$$

Desarrollando el lado derecho de la desigualdad se obtiene que

$$\limsup_{n\to\infty} \frac{1}{n} \log_{|R|} \frac{|R|^n}{V_w(n, (\delta n - 1)/2)}$$

puede expresarse como

$$\limsup_{n\to\infty} \frac{1}{n} \log_{|R|} |R|^n - \limsup_{n\to\infty} \frac{1}{n} \log_{|R|} V_w(n, (\delta n - 1)/2),$$

o de forma más simplificada como

$$1 - \limsup_{n\to\infty} \frac{1}{n} \log_{|R|} V_w(n, (\delta n - 1)/2).$$

Sea $h_w(\delta')$. Entonces, el argumento del volumen de la esfera deberá tener la forma

$$\frac{\delta n - 1}{2} = n\delta',$$

esto es,

$$\delta' = \frac{\delta - \frac{1}{n}}{2},$$

y puesto que $n \to \infty$

$$\delta' = \frac{\delta}{2}.$$

Por lo tanto,

$$\lim_{n\to\infty} \sup \frac{1}{n} \log_{|R|} \frac{|R|^n}{V_w(n, (\delta n - 1)/2)} = 1 - h_w\left(\frac{\delta}{2}\right)$$

y

$$\alpha_w(\delta) \leq 1 - h_w\left(\frac{\delta}{2}\right).$$

Para completar la demostración consideremos lo siguiente. Dado un código $C \subseteq R^n$ con distancia mínima d, si

$$\bigcup_{c\in C} B_w(c, d-1) \neq R^n$$

entonces C puede ser aumentado con una palabra tomada de $R^n \setminus \bigcup_{c\in C} B_w(c, d-1)$ y mantener si distancia mínima. Por esta razón

$$A_w(n, d) \cdot V_w(n, d-1) \geq |R|^n,$$

y aplicando primero el logaritmo de base $|R|$ y posteriormente el límite superior cuando $n \to \infty$ finalmente obtenemos

$$\alpha_w(\delta) \geq 1 - h_w(\delta)$$

completando así la prueba. □

El siguiente teorema proporciona una versión asintótica para la cota de Elias.

4.29 Teorema (Cota asintótica de Elias)**.** *Sea w un peso homogéneo con valor promedio γ sobre el anillo de Frobenius R. Entonces*

$$\alpha_w(\delta) \leq 1 - h_w\left(\gamma - \sqrt{\gamma(\gamma - \delta)}\right) \text{ si } 0 \leq \delta \leq \gamma;$$
$$\alpha_w(\delta) = 0 \text{ si } \gamma < \delta.$$

Demostración. Primero, observamos que $\alpha_w(\delta) = 0$ para $\gamma < \delta$ se sigue del Teorema 4.17 puesto que tenemos que

$$A_w(n, \delta n) \leq \frac{d}{d - \gamma n} = \frac{\delta}{\delta - \gamma}.$$

Puesto que la expresión anterior es constante en n, vemos que

$$\alpha_w(\delta) \leq \lim_{n\to\infty} \log_{|R|} \frac{\delta}{\delta-\gamma} = 0.$$

Para demostrar la primera parte del teorema, escribiremos la cota del Teorema 4.22 en la forma

$$A_w(n,\delta n) \leq \frac{\gamma\delta n^2}{t^2 - 2t\gamma n + \delta\gamma n^2} \frac{|R|^n}{V_{R_w}(n,t)},$$

siempre que $t \leq \gamma n$, y $t^2 - 2t\gamma n + \gamma\delta n^2 > 0$. Haciendo $t = \lambda n$:

$$A_w(n,\delta n) \leq \frac{\gamma\delta}{\lambda^2 - 2\lambda\gamma + \delta\gamma} \frac{|R|^n}{V_{R_w}(n,\lambda n)},$$

siempre que $\lambda \leq \gamma$ y $g(\lambda) := \lambda^2 - 2\lambda\gamma + \delta\gamma > 0$. De forma análoga al Ejemplo 4.23, encontramos que $g(\lambda) > 0$ para $\lambda_1 < \gamma - \sqrt{\gamma(\gamma-\delta)}$ y $\lambda_2 > \gamma + \sqrt{\gamma(\gamma-\delta)}$. Claramente, el valor de λ_2 se desprecia puesto que $\lambda \leq \gamma$. Por lo tanto, para cada $\lambda < \gamma - \sqrt{\gamma(\gamma-\delta)}$, podemos escribir

$$\frac{1}{n}\log_{|R|} A_w(n,\delta n) \leq \frac{1}{n}\log_{|R|} \frac{\gamma\delta}{\lambda^2 - 2\gamma\lambda + \gamma\delta} + 1 - \frac{1}{n}\log_{|R|} V_w(n,\lambda n).$$

El límite del segundo miembro de la desigualdad es, por definición, $1 - h_w(\lambda)$. Puesto que esto es cierto para cada $\lambda < \gamma - \sqrt{\gamma(\gamma-\delta)}$, se concluye que

$$\alpha_w(\delta) \leq 1 - h_w\left(\gamma - \sqrt{\gamma(\gamma-\delta)}\right).$$

□

4.30 Ejemplo. La Figura 4.1, misma que aparece al final de este capítulo, muestra las versiones asintóticas de las cotas de Gilbert-Varshamov:

$$1 - h_w(\delta) \leq \alpha_w(\delta),$$

de Elias:

$$\alpha_w(\delta) \leq 1 - h_w\left(\gamma - \sqrt{\gamma(\gamma-\delta)}\right),$$

y de Hamming:

$$\alpha_w(\delta) \leq 1 - h_w\left(\frac{\delta}{2}\right),$$

para el alfabeto $\mathbb{Z}_8$ equipado con $\gamma = 1$ y el peso homogéneo dado en el Ejemplo 4.23, donde la definición de $h_w(\delta)$ está dada como en el Ejemplo 4.27.

En la Tabla 4.2, también al final del capítulo, se resumen los valores de las respectivas cotas dados valores de δ. Note que se omiten de la tabla los valores $\delta = 0$ y $\delta = 1$, evitando así indeterminaciones.

4.31 Lema. *Para toda* $0 < x < \gamma$ *la función* $f(x) = \gamma - \sqrt{\gamma(\gamma - x)} - x/2$ *es positiva.*

Demostración. Evaluando la función $f(x)$ en $x = 0$ tenemos que $f(0) = \gamma - \gamma = 0$. Dado que la derivada de la función

$$f'(x) = \frac{\gamma}{2\sqrt{\gamma(\gamma - x)}} - \frac{1}{2} > 0$$

siempre que $\sqrt{\gamma(\gamma - x)} < \gamma$, se tiene que $f(x) > 0$ para $0 < x < \gamma$. □

4.32 Proposición. *La función* $h_w(\delta)$ *es una función creciente de* δ *y, por lo tanto, la cota asintótica de Elias es siempre más restrictiva que la cota asintótica de Hamming.*

Demostración. Deseamos probar que para $\delta_1 < \delta_2$

$$\min_{z \in (0,1]} \frac{f_w(z)}{z^{\delta_1}} \leq \min_{z \in (0,1]} \frac{f_w(z)}{z^{\delta_2}}.$$

Asumamos que, para $i \in \{1, 2\}$, $z_i \in (0, 1]$ es el valor en el que $f_w(z)/z^{\delta_i}$ alcanza su mínimo y definamos por simplicidad $F(z, \delta) := f_w(z)/z^{\delta}$. Demostraremos que $F(z_1, \delta_1) \leq F(z_2, \delta_1) \leq F(z_2, \delta_2)$. La primera desigualdad se sigue de z_1 minimizando $F(z, \delta_1)$. La segunda desigualdad se sigue de $F(z, \delta)$ siendo una función creciente en δ. Esto es cierto porque

$$\frac{\partial}{\partial \delta} F(z, \delta) = F(z, \delta)(-\ln z),$$

lo cual es no negativo para $z \in (0, 1]$. La proposición se sigue de combinar la última expresión con el lema precedente. □

δ	**Cota de Gilbert-Varshamov**	**Cota de Elias**	**Cota de Hamming**
0.04	0.88180	0.93356	0.93414
0.08	0.79108	0.87975	0.88180
0.12	0.71125	0.83039	0.83470
0.16	0.63884	0.78376	0.79108
0.20	0.57220	0.73907	0.75009
0.24	0.51040	0.69583	0.71125
0.28	0.45283	0.65375	0.67424
0.32	0.39909	0.61260	0.63884
0.36	0.34889	0.57220	0.60487
0.40	0.30204	0.53243	0.57220
0.44	0.25839	0.49319	0.54074
0.48	0.21787	0.45437	0.51040
0.52	0.18044	0.41592	0.48111
0.56	0.14610	0.37774	0.45283
0.60	0.11488	0.33980	0.42550
0.64	0.08687	0.30204	0.39909
0.68	0.06221	0.26440	0.37356
0.72	0.04108	0.22687	0.34889
0.76	0.02379	0.18942	0.32505
0.80	0.01073	0.15206	0.30204
0.84	0.00250	0.11488	0.27982
0.88	0.00006	0.07811	0.25839
0.92	0.00502	0.04246	0.23775
0.96	0.02088	0.01073	0.21787

Tab. 4.2: Tabla de cotas asintóticas para $\alpha_w(\delta)$.

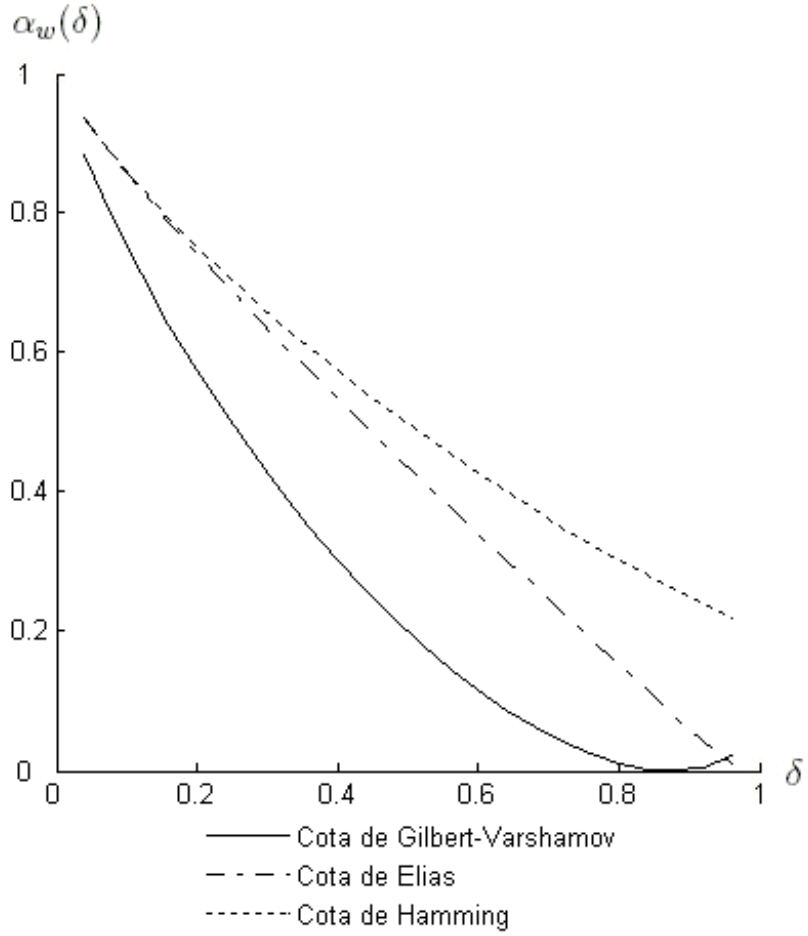


Fig. 4.1: Cotas asintóticas de Gilbert-Varshamov, Elias y Hamming.

CONCLUSIONES

A lo largo de este trabajo se ha hecho una revisión sobre los fundamentos de la teoría de códigos. Se establecieron las entidades involucradas en el proceso de transmisión de información y su papel en la teoría de códigos, así como los conceptos de código, distancia y peso. Asimismo, se analizaron diferentes métodos de decodificación (la decodificación por máxima similitud y la decodificación por distancias mínimas), la capacidad de un código para detectar y corregir errores y los posibles problemas a enfrentar que derivan en una decodificación de información incorrecta, así como las características de los códigos lineales y su codificación y decodificación (decodificación por el vecino más cercano y la decodificación por síndromes).

Se explicó que el problema principal de la teoría de códigos consiste en encontrar códigos óptimos para parámetros dados o, en otras palabras, en determinar el valor máximo de M para el que existe un código de longitud n, tamaño M y distancia minimal d sobre un alfabeto A de tamaño $q > 1$. En el marco de la teoría clásica de códigos se calcularon algunas de las cotas más sobresalientes: la cota de la esfera, la cota de Gilbert- Varshamov, la cota de Hamming (también conocida como "sphere packing bound"), la cota de Singleton, la cota de Plotkin y la cota de Griesmer; las primeras dos, cotas inferiores y las demás superiores. Además, se compararon estas cotas en el caso $A_2(n, 10)$ con $10 \leq n \leq 26$, siendo la de Plotkin la mejor de ellas, mientras que la de Hamming superó a la de Singleton de forma notable. Se calculó además una versión refinada de la cota de Plotkin para el caso de códigos binarios. A pesar de la elegancia y refinamiento propios de la cota de Plotkin, ésta posee limitaciones importantes ya que no es posible emplearla cuando $rn > d$, siendo $r = 1 - q^{-1}$, mientras que las cotas de Hamming y de Singleton son aplicables siempre que $q > 1$ y $1 \leq d \leq n$. El comparativo se realizó sin contemplar la cota de Griesmer en virtud de que ésta es válida únicamente sobre códigos lineales.

Por otra parte, se explicaron los elementos necesarios para interpretar el significado de caracter de un anillo y se abordó el concepto de caracter principal. Se estableció la definición de álgebra de Frobenius como una forma bilineal asociativa no degenerada y se explicó la construcción del caracter principal de un álgebra de Frobenius finita. La importancia de la caracterización de los anillos de Frobenius finitos radica en que un anillo finito R es de Frobenius si y sólo si R posee un caracter principal derecho (o, de forma

equivalente, un caracter principal izquierdo). Además, se presentó la relación que existe entre los anillos de Frobenius y el grupo cuántico de Lusztig. Se proporcionó también un algoritmo que permite reducir un producto arbitrario, y no necesariamente monótono, de elementos del conjunto de generadores W a una combinación lineal de palabras monótonas, lo cual permite a su vez calcular el caracter principal de un anillo.

Finalmente, se expuso que el desarrollo moderno de la teoría de códigos correctores de errores incluye la consideración de códigos sobre objetos más complejos que los campos finitos y que el problema principal en la teoría clásica de códigos prevalece cuando se contemplan códigos sobre anillos. Se estudió el caso de códigos sobre anillos finitos de Frobenius que cuentan con pesos homogéneos y se desarrollaron algunos resultados haciendo uso de la desigualdad de Cauchy-Schwarz. Se calcularon las cotas de Plotkin y de Elias para códigos sobre dichos anillos y se revisaron las versiones asintóticas de las cotas de Gilbert-Varshamov, de Hamming y de Elias. Se mostró además la razón por la cual la cota asintótica de Elias es más restrictiva que la cota asintótica de Hamming.

En el Capítulo 4 se estudió de forma detallada el artículo de Greferath y O'Sullivan (2004). Al hacer la revisión de dicho artículo se encontraron dos errores muy posiblemente de carácter tipográfico. Por otra parte, se encontró un resultado numérico que al parecer es incorrecto. En virtud de ello se escribió una carta a Marcus Greferath y Michael O'Sullivan, misma que se incluye en la sección de anexos. Hasta este momento O'Sullivan ha respondido que en cuanto les sea posible tanto Marcus como él harán una revisión del artículo respecto a las observaciones enviadas.

Para finalizar cabe mencionar que, en relación con los algoritmos, la computación cuántica abre posibilidades antes no imaginadas como son disminuciones exponenciales en el tiempo de procesamiento y la realización de operaciones en paralelo sin la necesidad de agregar procesadores a la máquina. A pesar de que en la actualidad no existe aún ninguna computadora cuántica real, resulta interesante proponer códigos sobre grupos cuánticos finitos para en un futuro estudiar la interrelación entre las áreas de cómputo y la teoría de códigos bajo el paradigma cuántico.

ANEXOS

A. Observaciones sobre artículo de M. Greferath y M. O'Sullivan

A.1. Carta a Marcus Greferath y Michael O'Sullivan

Universidad Nacional Autónoma de México
Centro de Investigaciones Teóricas
Av. 1 de Mayo, Col. Sta. María las Torres
Cuautitlán Izcalli, Estado de México
México, CP 54740.

September 11, 2008

Marcus Greferath and Michael O'Sullivan
San Diego State University
Department of Mathematics
5500 Campanile Drive, San Diego
CA 92182-7720, USA

Dear Professors Greferath and O'Sullivan,

I've been reading your article entitled *"On Bounds for Codes over Frobenius Rings under Homogeneous Weights"*, which has been very useful for my grade thesis.

Dr. Vladislav Khartchenko, my thesis adviser, and I have found two little mistakes in the cited article:

(i) On page 17, Corollary 3.3, $d\gamma n - 2W\gamma n + W^2 \geq 0$ must be strictly positive, i.e., $d\gamma n - 2W\gamma n + W^2 > 0$.

(ii) On page 18, after line 18, appears $M \leq \frac{|R^n|M_t}{V_w(n,t)}$ instead of $M \leq \frac{|R|^n M_t}{V_w(n,t)}$.

Perhaps both were typing errors when the document was being captured in LaTeX 2_ε.

Finally, on page 19 (after line 2) appears

$$\log_8(A_w(18,8)) \leq 11.72$$

.

According to Theorem 3.6, there are only five possible values for t: 0, 1, 2, 3 and 4. Considering that, for given positive integers $q > 1$, n and $t \geq 0$, $V_w(n,t)$ is defined as

$$V_w(n,t) = \begin{cases} \binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \ldots + \binom{n}{t}(q-1)^t & \text{if } 0 \leq t \leq n. \\ q^n & \text{if } n \leq t. \end{cases}$$

the Elias bound for each possible value of t is:

$$t = 0,\ \log_8(A_w(18,8)) \leq 18;$$
$$t = 1,\ \log_8(A_w(18,8)) \leq 15.80;$$
$$t = 2,\ \log_8(A_w(18,8)) \leq 14.01;$$

$$t = 3,\ \log_8(A_w(18,8)) \leq 12.51;$$

and

$$t = 4,\ \log_8(A_w(18,8)) \leq 11.44 = \log_8\left(\frac{144}{16 \cdot 7634572}\right) + 18,$$

which does not coincide with the exposed result on page 19.

I would appreciate if you could send me an explanation of this result.

Thank you very much.

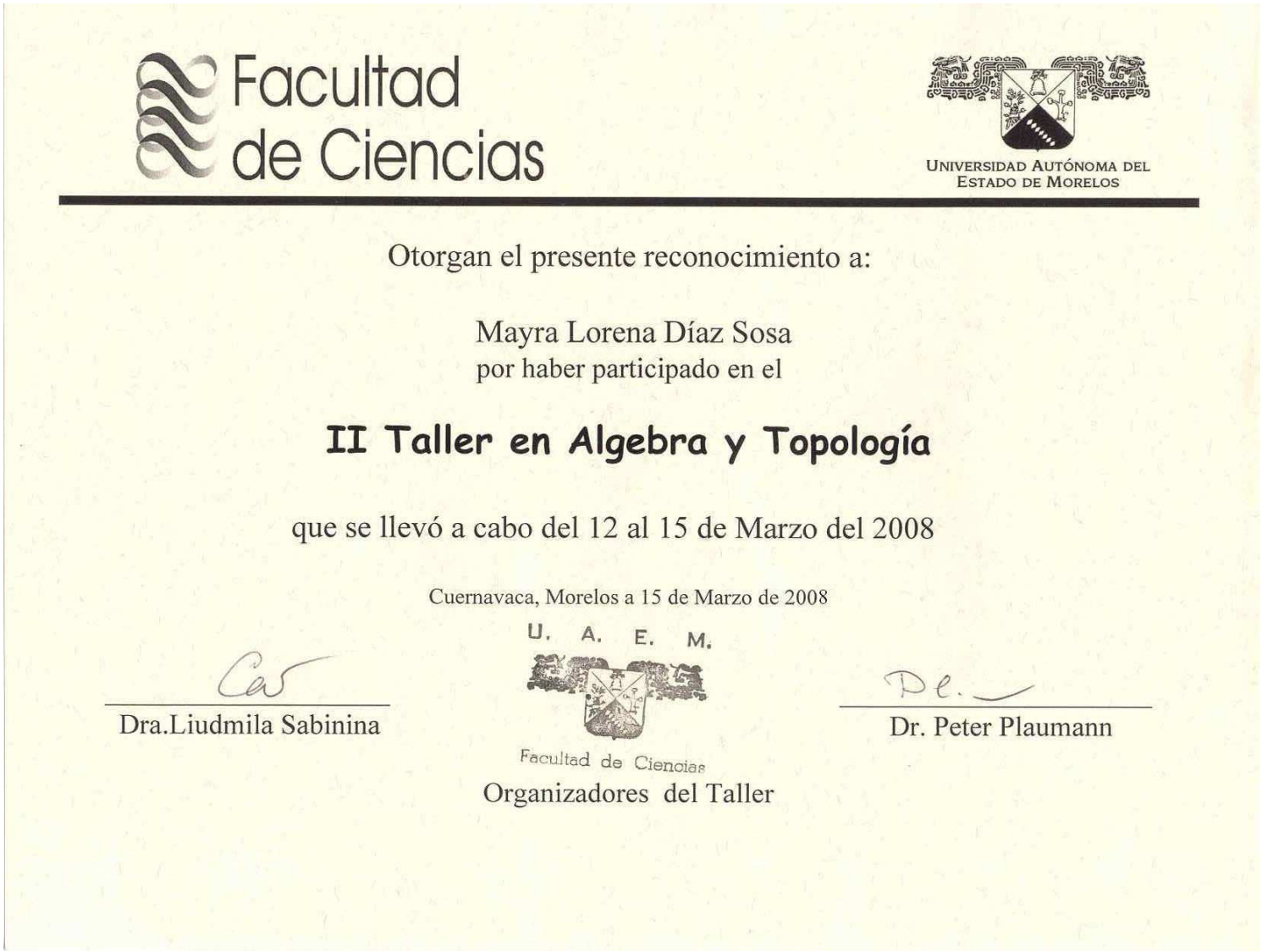
Yours sincerely,

Mayra Lorena Díaz Sosa
Computer Science Graduate Student
malodi1982@yahoo.com.mx

A.2. Respuesta de Michael O'Sullivan

```
----- Mensaje original ----
De: Michael E. O'Sullivan <mosulliv@sciences.sdsu.edu>
Para: Mayra Lorena Díaz Sosa <malodi1982@yahoo.com.mx>
CC: greferath@math.sdsu.edu; mosulliv@math.sdsu.edu;
    vlad@servidor.unam.mx
Enviado: miércoles, 17 de septiembre, 2008 9:22:55
Asunto: Greetings from Mexico and some remarks about your article

   Dear Mayra,

   Thank you for your note and careful reading of our paper.
   Marcus and I have some pressing matters to take care of, but we will
   take a look at your comments soon.

   All the best,

   Mike

   ------------------------------------------------
   Michael E. O'Sullivan
   Dept. of Mathematics  and Statistics
   San Diego State University
   5500 Campanile Dr.
   San Diego CA 92182-7720

   http://www-rohan.sdsu.edu/~mosulliv
   mosulliv@math.sdsu.edu
   GMCS 579
   (619) 594-6697 (my office)
   (619) 594-6191 (dept. office)
   (619) 594-6746 (FAX)
```

B. Constancias



Facultad de Ciencias

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

Otorgan el presente reconocimiento a:

Mayra Lorena Díaz Sosa
por haber participado en el

II Taller en Algebra y Topología

que se llevó a cabo del 12 al 15 de Marzo del 2008

Cuernavaca, Morelos a 15 de Marzo de 2008

U. A. E. M.

Dra.Liudmila Sabinina

Dr. Peter Plaumann

Facultad de Ciencias

Organizadores del Taller

Centro de Investigación en Matemáticas A.C.

Escuela de Modelación y Métodos Numéricos

CIMAT

Otorga la presente:

CONSTANCIA A:

Mayra Lorena Díaz Sosa

Por su asistencia a la
"Escuela de Modelación y Métodos Numéricos" Supercómputo y Aplicaciones
del 18 al 21 de junio en las instalaciones de este centro

Guanajuato, Gto junio del 2008

Dr. Salvador Botello Rionda
Comité Organizador

Dr. Miguel Ángel Moreles Vázquez
Comité Organizador

REFERENCIAS

Adamek, J. (1991). *Foundations of Coding: Theory and Applications of Error-Correcting Codes with an Introduction to Cryptography and Information Theory*. J. Wiley, Chichester.

Aerts, D. y M., Czachor (2008). «Tensor-product vs. geometric-product coding». *Physical Review A*, **77**, pp. 312–316.

Agrell, E.; Vardy, A. y Zeger, K. (2001). «A table of upper bounds for binary codes». *Information Theory, IEEE Transactions on*, **47(7)**, pp. 3004–3006.

Berlekamp, E. (1968). *Algebraic Coding Theory*. Mcgraw-hill series in systems science. McGraw-Hill, New York.

—— (1973). «Goppa codes». *Information Theory, IEEE Transactions on*, **19(5)**, pp. 590–592.

—— (1974). *Key Papers in the Development of Coding Theory*. IEEE press selected reprint series. Institute of Electrical and Electronics Engineers, New York.

Best, M. (1980). «Binary codes with a minimum distance of four». *Information Theory, IEEE Transactions on*, **26(6)**, pp. 738–742.

Best, M.; Brouwer, A.; MacWilliams, F.; Odlyzko, A. y Sloane, N. (1978). «Bounds for binary codes of length less than 25». *Information Theory, IEEE Transactions on*, **24(1)**, pp. 81–93.

Beth, T.; Egner, S.; Geiselmann, W.; Lazic, D. y Müller-Quade, J. (2003). *Computer Algebra Handbook*. Springer.

Bierbrauer, J. (2005). *Introduction to Coding Theory*. Discrete mathematics and its applications. Chapman & Hall/CRC, Boca Raton, Florida.

Birkoff, G. y Mac Lane, S. (1965). *A Survey of Modern Algebra*. Macmillan.

Brouwer, A. E. (2004). «Coding Theory». Disponible en línea en `http://www.win.tue.nl/~aeb/`. Último acceso en 11-08-2008.

Brouwer, A. E. y Verhoeff, T. (1993). «An updated table of minimum-distance bounds for binary linear codes». *Information Theory, IEEE Transactions on*, **39(2)**, pp. 662–677.

Byrne, E.; Greferath, M. y E., O'Sullivan M. (2007a). «Errata for "The linear programming bound for codes over finite Frobenius rings"». *Des. Codes Cryptogr.*, **45**, pp. 269–270.

—— (2007b). «The linear programming bound for codes over finite Frobenius rings». *Des. Codes Cryptogr.*, **42**, pp. 289–301.

Calderbank, A. R. (1998). «The art of signaling: fifty years of coding theory». *Information Theory, IEEE Transactions on*, **44(6)**, pp. 2561–2595.

Calderbank, A. R.; Hammons Jr, A. R.; Kumar, P. V.; Sloane, N. J. A. y Sole, P. (1993). «A Linear Construction for Certain Kerdock and Preparata Codes». *American Mathematical Society*, **29(2)**, pp. 218–222.

Cohen, G.; Karpovsky, M.; Mattson, H. y Schatz, J. (1985). «Covering radius-Survey and recent results». *Information Theory, IEEE Transactions on*, **31(3)**, pp. 328–343.

Constantinescu, I. y Heise, W. (1997). «A Metric for Codes over Residue Class Rings». *Problemy Peredachi Informatsii*, **33(3)**, pp. 22–28.

Conway, J. H. y Sloane, N. J. A. (1990). «A new upper bound on the minimal distance of self-dual codes». *Information Theory, IEEE Transactions on*, **36(6)**, pp. 1319–1333.

—— (1993). «Self-dual codes over the integers modulo 4». *J. Combin. Theory Ser. A.*, **62(1)**, pp. 30–45.

Dascalescu, S.; Nastasescu, C. y Raianu, S. (2001). *Hopf Algebras, An Introduction.* Marcel Dekker.

Dinh, H. Q. y López-Permouth, S. R. (2004a). «Cyclic and negacyclic codes over finite chain rings». *Information Theory, IEEE Transactions on*, **50(8)**, pp. 1728–1744.

—— (2004b). «On the Equivalence of Codes over Finite Rings». *Appliable Algebra in Engineering, Communication and Computing*, **15**, pp. 37–50.

—— (2004c). «On the equivalence of codes over rings and modules». *Finite Fields and Their Applications*, **10**, pp. 615–625.

Duursma, I. M.; Greferath, M.; Litsyn, S.Ñ. y Schmidt, S. E. (2001a). «A $\mathbb{Z}_8$ linear lift of the binary Golay code and a nonlinear binary (96, 2^{37}, 24)-code». *Information Theory, IEEE Transactions on*, **47(4)**, pp. 1596–1598.

—— (2001b). «A $\mathbb{Z}_9$ linear lift of the ternary [24, 12, 9]-code inducing a nonlinear ternary (72, 3^{25}, 24)-code». *Proceedings of optimal codes (OC 2001), Slantchev Briag, Bulgaria*, pp. 59–64.

Gerzson, K. (2008). «Tables for Bounds on Covering Codes». Disponible en línea en `http://www.sztaki.hu/~keri/codes/`. Último acceso en 11-08-2008.

Goldwasser, S. (2002). «Mathematical Foundations of Modern Cryptography: Computational Complexity Perspective». En: *ICM '02: Proceedings of the International Congress of Mathematicians*, Higher Education Press, Beijing, China.

Grassl, M. (2007). «Bounds on the minimum distance of linear codes». Disponible en línea en `http://www.codetables.de`. Último acceso en 11-08-2008.

Greferath, M.; McGuire, G. y O'Sullivan, M. E. (2006). «On Plotkin optimal codes over finite Frobenius rings». *Journal of Algebra and Its Applications*, **5**, pp. 799–815.

Greferath, M. y O'Sullivan, M. E. (2004). «On bounds for codes over Frobenius rings under homogeneous weights». *Discrete Mathematics*, **289**, pp. 11–24.

Greferath, M. y Schmidt, S. E. (2000). «Finite-Ring Combinatorics and MacWilliams' Equivalence Theorem». *Journal of Combinatorial Theory, Series A*, **92(1)**, pp. 17–28.

Grimaldi, R. P. (1994). *Discrete and Combinatorial Mathematics, An Applied Introduction*. Addison Wesley, 3ª edición.

Haas, W. (2008). «On the failing cases of the Johnson bound for error correcting codes». *The Electronic Journal of Combinatorics*, **15(R55)**, p. 1.

Haaser, N. B.; LaSalle, J. P. y Sullivan, J. A. (1972). *Análisis matemático 1: Curso de introducción*. Trillas.

Hammons Jr., A. R.; Kumar, P. V.; Calderbank, A. R.; Sloane, N. J. A.; Sole, P.; Div, N. S.; Co, H. A. y Germantown, M. D. (1994). «The $\mathbb{Z}_4$ linearity of Kerdock, Preparata, Goethals, and related codes». *Information Theory, IEEE Transactions on*, **40(2)**, pp. 301–319.

Hartnett, W. E. (1974). *Foundations of Coding Theory*. volumen 1 de *Episteme*. Holland d. reidel, Dordrecht.

Herstein, I. (1975). *Topics in Algebra*. John Wiley.

—— (1996). *Noncommutative Rings*. Número 15 en The Carus Mathematical Monographs. The Mathematical Associaton of America, 4ª edición.

Honold, T. (2001). «Characterization of finite Frobenius rings». *Archiv der Mathematik*, **76**, pp. 406–415.

Janwa, H. (1989). «Some new upper bounds on the covering radius of binary linear codes». *Information Theory, IEEE Transactions on*, **35(1)**, pp. 110–122.

Johnson, S. (1962). «A new upper bound for error-correcting codes». *Information Theory, IEEE Transactions on*, **8(3)**, pp. 203–207.

Joseph, A. (1994). *Quantum Groups and Their Primitive Ideals*. Springer-Verlag.

Klemm, M. (1989). «Selbstduale Codes über dem Ring der ganzen Zahlen modulo 4». *Archiv der Mathematik*, **53(2)**, pp. 201–207.

Kustermans, J. (1997). «Examining the dual of an algebraic quantum group». *eprint arXiv: funct-an/9704004*.

Kuzmin, A. S.; Markov, V. T.; Nechaev, A. A. y Neljubin, A. S. (2006). «A generalization of the binary Preparata code». *Discrete Applied Mathematics*, **154**, pp. 337–345.

Larson, R. G. y Sweedler, M. (1969). «An associative orthogonal bilinear form for Hopf algebras». *American Journal of Mathematics*, **91**, pp. 75–93.

Lay, D. C. (1997). *Linear algebra and its applications*. Addison-Wesley Reading, Mass.

Ling, S. y Xing, Ch. (2004). *Coding Theory, A First Course*. Cambridge University Press.

Loeliger, H. A. (1994). «An upper bound on the volume of discrete spheres». *Information Theory, IEEE Transactions on*, **40(6)**, pp. 2071–2073.

Lusztig, G. (1994). *Introduction to Quantum Groups*. Birkhäuser.

Mounits, B.; Etzion, T. y Litsyn, S. (2002). «Improved upper bounds on sizes of codes». *Information Theory, IEEE Transactions on*, **48(4)**, pp. 880–886.

Mueger, M.; Roberts, J. E. y Tuset, L. (2004). «Representations of algebraic quantum groups and reconstruction theorems for tensor categories». *Algebras and Representation Theory*, **7**, p. 517.

Nechaev, A. A.; Kuzmin, A. S. y Markov, V. T. (1997). «Linear codes over finite rings and modules». *Fundamentalnaya i Prikladnaya Matematika*, **3(1)**, pp. 195–254.

Ostergard, P. R. J.; Baicheva, T. y Kolev, E. (1999). «Optimal binary one-error-correcting codes of length 10 have 72 codewords». *Information Theory, IEEE Transactions on*, **45(4)**, pp. 1229–1231.

Pless, V. y Qian, Z. (1996). «Cyclic codes and quadratic residue codes over $\mathbb{Z}_4$». *Information Theory, IEEE Transactions on*, **42(5)**, pp. 1594–1600.

Pless, V.; Solé, P. y Qian, Z. (1997). «Cyclic Self-Dual $\mathbb{Z}_4$-Codes». *Finite Fields and Their Applications*, **3(1)**, pp. 48–69.

Plotkin, M. (1960). «Binary codes with specified minimum distance». *Information Theory, IEEE Transactions on*, **6(4)**, pp. 445–450.

Pretzel, O. (1998). *Error-Correcting Codes and Finite Fields*. Oxford Applied Mathematics and Computing Science Series. Clarendon Press.

Quistorff, J. (2003). «Some remarks on the Plotkin bound». *The Electronic Journal of Combinatorics*, **10(6)**, p. 1.

Ritter, W. G. (2002). «Introduction to Quantum Group Theory». *Arxiv preprint math/0201080*.

Robinson, D. J. (1996). *A Course in the Theory of Groups*. Springer-Verlag, 2ª edición.

Shannon, C. E. (1948). «A mathematical theory of communication». *Bell System Technical Journal*, **27**, pp. 379–423, 623–656.

—— (2001). «A mathematical theory of communication». *ACM SIGMOBILE Mobile Computing and Communications Review*, **5(1)**, pp. 3–55.

Singleton, R. (1964). «Maximum distance q-nary codes». *Information Theory, IEEE Transactions on*, **10(2)**, pp. 116–118.

Skryabin, S. (2007). «Projectivity and freeness over comodule algebras». *Transactions of the American Mathematical Society*, **359(6)**, pp. 2597–2623.

Sudan, M. (2001). «Coding Theory: Tutorial and Survey», p. 36.

Tietavainen, A. (1973). «On the nonexistence of perfect codes over finite fields». *SIAM Journal on Applied Mathematics*, **24(1)**, pp. 88–96.

—— (1990). «An upper bound on the covering radius as a function of the dual distance». *Information Theory, IEEE Transactions on*, **36(6)**, pp. 1472–1474.

van Lint, J. H. (1975). «A survey of perfect codes». *Rocky Mountain J. Math*, **5(2)**, pp. 199–226.

—— (1999). *Introduction to Coding Theory*. Springer, 3ª edición.

Vu, V. y Wu, L. (2005). «Improving the Gilbert–Varshamov Bound for q-ary Codes». *Information Theory, IEEE Transactions on*, **51(9)**, pp. 3200–3208.

Ward, H. y Wood, J. A. (1996). «Characters and the equivalence of codes». *Journal of Combinatorial Theory, Series A*, **73(2)**, pp. 348–352.

Wood, J. A.. «Extension Theorems for Linear Codes over Finite Rings».

—— (1999). «Duality for modules over finite rings and applications to coding theory». *American Journal of Mathematics*, **121**, pp. 555–575.

—— (2006). «A coding-theoretic characterization of Finite Frobenius Rings».