Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
LICENCIATURA EN MATEMÁTICAS

EL ABC DE LOS PRELIMINARES BÁSICOS PARA EL ESTUDIO DE LOS CÓDIGOS ALGEBRAICOS.

TESIS

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS

PRESENTA
Luis Felipe Munguia Aca

DIRECTORES DE TESIS
Dr. Iván Fernando Vilchis Montalvo

PUEBLA, PUE. FECHA: FEBRERO 2019

Dedicatoria

A mis padres.

Por haberme apoyado en todos estos años de la carrera, por confiar en mí, y por siempre estar al pendiente de mí.

A mi maestros.

Dr. Iván Fernando Vilchis Montalvo por haberme aceptado como su alumno, por su apoyo, su amistad, sus consejos, y por la inspiración a seguir en el área de las matemáticas, al Dr. Francisco Javier Mendoza Torres, por su gran apoyo y motivación para la culminación de mis estudios, al Dr. Luis Alberto Guerrero Méndez, por ser uno de mis mejores amigos y maestros de la carrera.

A mi novia.

Alely Maldonado Azcona por brindarme en estos últimos meses su compañia, cariño y motivación.

A mis amigos.

Que nos apoyamos mutuamente en nuestra formación personal dentro y fuera del aula y que hasta ahora los considero más que mis amigos mis hermanos: Daniel Fuentes, Daniel Hernández , Hector Manuel López, Luis Enrique Luna, Pedro Teutle, Reymundo Guerrero.

Agradecimientos

Esté trabajo de tesis fue logrado gracias a la participación de varias personas, a quienes les estaré eternamente agradecido, ya que cada una de estás personas han puesto su granito de arena.
Agradezco principalmente a mi asesor de tesis el Dr. Iván Fernando Vilchis Montalvo, por brindarme su tiempo, paciencia, su apoyo, y conocimientos, así como brindarme un agradable y acojedor lugar de trabajo.
Agradezco a mi jurado el Dr. Carlos Alberto López Andrade, el Dr. David Villa Hernández y el Dr. César Cejudo Castilla, por tomarse el tiempo, la molestia de revisar, corregir y también por darme sus valiosas observaciones.
Agradezco a mi amigo Daniel Fuentes, por su ayuda brindada en partes especificas de este trabajo.

Introducción

El estudio de los anillos finitos conmutativos, además de tener un cierto intéres en el aspecto teórico tiene bastantes aplicaciones, una de las cuales es la teoría de la codificación. En este trabajo nuestra intención es dar una introducción sutil a los temas de la teoría de códigos y de los anillos conmutativos finitos.
Así, por lo anteriormente dicho, estás notas hablarán sobre la teoría de códigos. En el día a día convivimos con muchos códigos aunque no nos demos cuenta. Por ejemplo, los más comunes son el código de barras, el ISBN usado en los libros y el código ASCII usado en las computadoras. Quizá los primeros códigos usados son el código Morse, usado en la telegrafía desde el siglo XIX, y el sistema Braile para no-videntes. Además cualquier artefacto tecnológico, que transmita o almacene mensajes, imagenes o sonidos, involucra al menos un código. Por mencionar a algunos, tenemos las computadoras, celulares, satelites, C.D's, D.V.D's, etc.
La situación en la que nos vamos a encontrar es la que muestra la Figura 1: Supongamos que queremos enviar un mensaje **x**. La idea es que, antes de enviarlo, codifiquemos el mensaje **x** como **c**. Lo común, es reescribir el mensaje (en el caso no trivial) en forma diferente incluso usando un alfabeto distinto, pero nosotros lo haremos bajo determinadas reglas. Una de ellas es que cada mensaje (palabra) no puede tener más de una palabra código en el código. Además debemos añadirle a **x** información redundante, de tal forma que si en el canal de transmisión se produce ruido **r** y el receptor en vez de **c** recibe un mensaje alterado **c'=c+r** sea, a pesar de todo, capaz de recuperar el mensaje original y si no al menos el más probable.

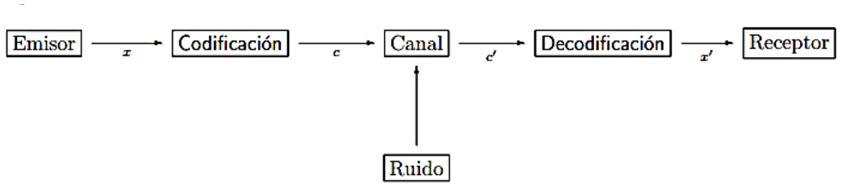

Figura 1: Figura 1

Supongamos la siguiente situación concreta: Imaginemos que somos los

encargados en el cambio de vías de un tren, que manejamos a control remoto desde nuestra cabina por medio de un canal que transmite impulsos eléctricos de dos voltajes distintos, que denotaremos por 0 y 1 respectivamente. El cambio de vías puede hacer que el tren siga su curso normal, gire a la derecha, a la izquierda o incluso regresar por donde vino. Luego, nuestros mensajes son N, R, D e I. Y los codificamos por ejemplo:

$$N \to 10,\ R \to 01,\ D \to 00,\ I \to 11$$

Ahora supongamos que necesitamos que el tren tome la vía de la derecha. Enviamos el mensaje 00. Si ocurre una interferencia en la transmisión hace que el cambio de vías reciba 01, es decir que regrese. El problema está en nuestro código

$$C = \{00, 01, 10, 11\}$$

que no detecta errores. Ya que, si hay un error en la transmisión, la palabra recibida es otra palabra código. Para tratar de arreglar esto, podemos agregar redundancia por ejemplo agregando un dígito extra a cada palabra código de modo que la suma de los dígitos de cada palabra código sea 0 o par. Así el nuevo código será el siguiente:

$$C' = \{000, 011, 101, 110\} \subset \mathbb{Z}_2^3$$

Por lo tanto si enviamos el mensaje 000 y se recibe digamos 010. Como la palabra código no pertenece al código, se detecta un error, no se produce el cambio de vía, y por lo tanto podríamos volver a intentar mandar el mensaje.

Índice general

Título de la tesis

Autor

fecha

Capítulo 1

Algunos resultados básicos de álgebra

Primero vamos a recordar algunos conceptos básicos del álgebra, en particular la teoría de los anillos conmutativos finitos que es la base para el desarrollo de la teoría de códigos.
De ahora en adelante, consideraremos a R como un anillo conmutativo con unidad.

Definición 1.1. *Un subconjunto S de un anillo R es un* ***subanillo*** *de R si*

1. $1 \in S$.
2. *Si* $a, b \in S$, *entonces* $a - b \in S$.
3. *Si* $a, b \in S$, *entonces* $ab \in S$.

Ejemplo 1.2. *El conjunto*

$$S = \{ \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \mid a, b, c \in \mathbb{Z} \}$$

es un subanillo del anillo $\mathbb{M}_{2\times 2}(\mathbb{Z})$.

Definición 1.3. *Un elemento a en un anillo R es llamado un* ***divisor de cero*** *si $a \neq 0$ y si existe $0 \neq b \in R$ tal que $ab = 0$.*

Ejemplo 1.4. *En el anillo $\mathbb{Z}_6$, los elementos $\overline{3}$ y $\overline{4}$ son distintos de la clase $\overline{0}$, pero $(\overline{3})(\overline{4}) = \overline{12} = \overline{0}$ en $\mathbb{Z}_6$, por lo tanto $\overline{3}$ y $\overline{4}$ son divisores de cero.*

Definición 1.5. *Sea $R \neq \{0\}$ un anillo. Decimos que R es un* ***dominio entero*** *si no contiene divisores de cero no triviales.*

Ejemplo 1.6. *$(\mathbb{Z}, +, \cdot)$*

Definición 1.7. *Un elemento $0 \neq x \in R$ es* ***nilpotente*** *si $x^n = 0$ para algún entero positivo n.*

Ejemplo 1.8. *1. En el anillo $\mathbb{Z}_9$, el elemento $\overline{3}$ es nilpotente, ya que $\overline{3}^2 = \overline{3^2} = \overline{9} = \overline{0}$.*

2. Un elemento nilpotente es un divisor de cero en R, siempre que R no sea el anillo trivial, es decir, $R = 0$.

Definición 1.9. *Sea $u \in R$, decimos que u es unidad si exite $v \in R$ tal que $uv = 1$, denotamos $v = u^{-1}$.*

Ejemplo 1.10. *En el anillo de los enteros, $\mathbb{Z}$, las unidades son precisamente 1 y -1.*

Definición 1.11. *El subconjunto de R*

$$U(R) := \{x \in R \mid \exists\ y \in R \text{ tal que } xy = yx = 1\}$$

es un grupo multiplicativo con respecto al producto en R, y sus elementos son llamados las unidades de R.

Ejemplo 1.12. *En $\mathbb{Z}$, $U(\mathbb{Z}) = \{1, -1\} \cong \mathbb{Z}_2$. Donde el isomorfismo es de grupos abelianos.*

Definición 1.13. *Un anillo $R \neq \{0\}$ es un* ***campo*** *si todo elemento no cero es unidad. Es decir, $U(R) = R^* := R \setminus \{0\}$.*

Para este trabajo denotaremos a los campos por $\mathbb{F}$. Diremos que $\mathbb{F}_q$ es un campo finito si tiene un número q de elementos, donde $q \in \mathbb{N}$.

Ejemplo 1.14. *Los ejemplos más comunes de campos infinitos son $\mathbb{Q}, \mathbb{R}$ y $\mathbb{C}$ y el ejemplo más común de campo finito son los enteros $\mathbb{Z}$ mod p denotado por $\mathbb{Z}_p$.*

Definición 1.15. *Un* ***subcampo*** *de un campo $\mathbb{F}$ es un subanillo k de $\mathbb{F}$ que es también un campo.*

Proposición 1.16. *Todo campo es un dominio entero.*

Demostración. Ya que $ab = 0$ y $a \neq 0$ implica que $b = 1b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0$. □

Proposición 1.17. *Todo dominio entero finito es un campo.*

Demostración. Asumamos que $\{a_1, a_2, ..., a_n\}$ son los elementos de un dominio entero finito R. Para un elemento $0 \neq a \in R$ consideremos todos los productos $\{aa_1, aa_2, ..., aa_n\}$. Estos son distintos, porque si $aa_i = aa_j$, entonces $a(a_i - a_j) = 0$, y ya que $a \neq 0$, tenemos que $a_i = a_j$. Se sigue que cada elemento de R es de la forma aa_j; en particular, existe $h \in \{1, ..., n\}$ tal que $1_R = aa_h$. Ya que R es conmutativo, tenemos también $1_R = a_h a$, entonces $a_h = a^{-1}$. Por lo tanto, los elementos no cero de R forman un grupo abeliano con respecto a la multiplicación. □

Ejemplo 1.18. *Son equivalentes las siguientes condiciones:*

I. $\mathbb{Z}_p$ *es campo.*

II. $\mathbb{Z}_p$ *es dominio entero.*

III. p *es primo.*

Demostración. i. $\Rightarrow$] ii. Por la Proposición 1.16.
ii. $\Rightarrow$] i. Por la Proposición 1.17.
ii. $\Rightarrow$] iii. Si p no es primo, entonces existen $a, b \in \mathbb{Z}$, con $1 < a, b < p$, tales que $p = ab$. Luego $\overline{ab} = \overline{p}$ en $\mathbb{Z}_p$ implica que $\overline{a}\overline{b} = \overline{p} = \overline{0}$, por lo tanto $\mathbb{Z}_p$ no es dominio entero, contradicción.
iii. $\Rightarrow$] ii. Sean $\overline{a}$,$\overline{b}$ en $\mathbb{Z}_p$ y p primo. Supongamos que $\overline{ab} = \overline{0}$, entonces $p \mid ab$, por lo que $p \mid a$ ó $p \mid b$, pero esto significa que $\overline{a} = \overline{0}$ ó $\overline{b} = \overline{0}$ en $\mathbb{Z}_p$, por lo que $\mathbb{Z}_p$ es dominio entero. □

Lema 1.19. *Sean a, b dos elementos de un campo $\mathbb{F}$ entonces,*

I. $(-1)a = -a$.

II. $ab = 0$ *implica* $a = 0$ *ó* $b = 0$.

Demostración. I. $(-1)a + a = (-1)a + 1a = ((-1) + 1)a = 0a = 0$. Por lo tanto $(-1)a = -a$, por la unicidad de los inversos.

II. Si $a \neq 0$, entonces $0 = a^{-1}0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$.

□

Nota: La condición i. se cumple para cualquier anillo y la ii. se cumple par cualquier dominio entero.

Definición 1.20. *Para un anillo R, un entero $n \geq 1$ y $a \in R$ denotamos por na el elemento,*

$$\sum_{k=1}^{n} a = a+a+a+...+a$$

Definición 1.21. *Sea $\mathbb{F}$ un campo, la característica de $\mathbb{F}$ es el menor entero positivo p tal que $p1 = 0$, donde 1 es el neutro multiplicativo de $\mathbb{F}$. Si tal p no existe decimos que es de característica 0.*

Proposición 1.22. *La característica de un campo es 0 o un número primo.*

Demostración. Es claro que 1 no puede ser la característica, ya que $1(1) = 1 \neq 0$. Supongamos que la característica del campo $\mathbb{F}$ es p con $p = nm$ con $1 < n, m < p$, pero

$$(n1)(m1)=(\sum_{i=1}^{n} 1)(\sum_{j=1}^{m} 1)=(mn)1=p1 =0.$$

por el Lema 1.19, $m1 = 0$ ó $n1$=0. Contradicción.

□

Proposición 1.23. *Si $\mathbb{F}$ es un campo de característica $p \geq 0$, entonces $pa = 0$ para todo $a \in \mathbb{F}$.*

Demostración. Ya que $\mathbb{F}$ tiene característica p, tenemos que $p1 = 0$, donde 1 es el uno de $\mathbb{F}$. Así

$$pa = p(1a) = (p1)a = 0a = 0.$$

□

Proposición 1.24. *Sea $\mathbb{F}$ un campo y p primo, entonces p divide $\binom{p}{r}$ para todo $1 \leq r \leq p-1$.*

Demostración. Para p primo, $\binom{p}{r} = \frac{p!}{r!(p-r)!}$. Ya que $\binom{p}{r}$ es un entero, $r!(p-r)!$ divide $p!$ para $1 \leq r \leq p-1$. Como $\binom{p}{r} = \frac{p}{r}\binom{p-1}{r-1}$, seguidamente multiplicando por r en ambos lados, tenemos $r\binom{p}{r} = p\binom{p-1}{r-1}$, así $p \mid r\binom{p}{r}$, entonces $p \mid r$ ó $p \mid \binom{p}{r}$. Por lo tanto $p \mid \binom{p}{r}$. □

Proposición 1.25. *Si $\mathbb{F}$ es un campo de característica p y $a, b \in \mathbb{F}$, entonces $(a+b)^{p^n} = a^{p^n} + b^{p^n}$.*

Demostración. Por el Teorema Binomial

$$(a+b)^p = a^p + \binom{p}{1}a^{p-1}b + \binom{p}{2}a^{p-2}b^2 + ... + \binom{p}{p-1}ab^{p-1} + \binom{n}{n}b^n$$

Usando las Proposiciones 1.24 y 1.23 concluimos $(a+b)^p = a^p + b^p$. Haciendo inducción sobre n.

1. Cuando $n = 1$, entonces $(a+b)^p = a^p + b^p$.
2. Supongamos que se cumple para $n = k$, es decir
$$(a+b)^{p^k} = a^{p^k} + b^{p^k}$$
3. Veamos que se cumple para $n = k+1$.
$$\begin{aligned}(a+b)^{p^{k+1}} &= (a+b)^{p^k p} \\ &= ((a+b)^{p^k})^p \\ &= (a^{p^k} + b^{p^k})^p \\ &= (a^{p^k})^p + (b^{p^k})^p \\ &= a^{p^{k+1}} + b^{p^{k+1}}.\end{aligned}$$

□

Proposición 1.26. *Sea $p \in \mathbb{N}$ primo. Un campo finito $\mathbb{F}$ de característica p contiene p^n elementos para algún natural $n \geq 1$.*

Demostración. Sea p la caracteristica de $\mathbb{F}$. Como $1 \in \mathbb{F}$, tenemos que $p1 = 0$, y dado que $\mathbb{F}$ es finito, entonces el orden de 1 es igual a p y es tal que $p \mid |\mathbb{F}|$. Supongamos que existe $q \neq p$ primo tal que $q \mid |\mathbb{F}|$. Por el Teorema de Cauchy existe $x \in \mathbb{F}$ tal que el orden de x es q, así $m.c.d(p,q) = 1$, lo que implica que existen $s, t \in \mathbb{Z}$ tales que $sp + tq = 1$, multiplicando por x en ambos lados tenemos que $x = s(px) + t(qx) = 0$, lo que es una contradicción. Por lo tanto $|\mathbb{F}| = p^n$. □

Capítulo 2

Anillo de Polinomios

En la teoría de anillos conmutativos, el anillo de polinomios sobre un anillo R es quizás el anillo más importante de está teoría debido a su interdependencia con los demás conceptos. En particular en este trabajo su importancia radicará en el estudio de los códigos cíclicos y en algunos resultados sobre el anillo de polinomios sobre los anillos de Galois.

Definición 2.1. *Sea R un anillo. Una sucesión en R es una función*

$$\begin{aligned} \sigma : \mathbb{N} &\longrightarrow R \\ i &\longmapsto s_i \end{aligned}$$

Notamos que σ la podemos representar de la siguiente manera:

$$\sigma = (s_0, s_1, ..., s_i, ...).$$

*A las entradas $s_i \in R$, para todo $i \in \mathbb{N}$, los llamamos los **coeficientes** de σ.*

Sea τ: $\mathbb{N} \longrightarrow R$ y $\sigma : \mathbb{N} \longrightarrow R$, donde $\tau = (t_0, ..., t_i, ...)$ y $\sigma = (s_0, ..., s_i, ...)$. Decimos que $\tau = \sigma$ sí y sólo si $s_i = t_i$ para toda $i \in \mathbb{N}$.

Definición 2.2. *Una sucesión $\sigma = (s_0, s_1, ..., s_i, ...)$ en R es llamado un **polinomio** si hay algún natural $m \geq 0$ con $s_i = 0$, para todo $i > m$, esto es,*

$$\sigma = (s_0, s_1, ..., s_m, 0, 0, ...).$$

Un polinomio tiene a lo más un número finito de coeficientes distintos de cero. El polinomio cero, denotado por $\sigma = \mathbf{0}$ es la sucesión $\sigma = (0, 0, ..., 0)$.

Definición 2.3. *Si $0 \neq \sigma = (s_0, s_1, ..., s_m, 0, 0, ...)$ es un polinomio, entonces existe $s_n \neq 0$ con $s_i = 0$ para todo $i > n$. Decimos que s_n es el* ***coeficiente principal*** *de σ y que n es el* ***grado*** *de σ, denotado por* ***gr***$(\sigma) = n$.

El polinomio cero 0 no tiene un grado porque no tiene coeficientes distintos de cero. Sera conveniente decir que el grado del polinomio cero es el simbolo $-\infty$ y adoptar las convenciones habituales que $-\infty < n$ para todo $n \in \mathbb{N}$, $-\infty + (-\infty) = \infty$, $-\infty + n = \infty$.

Notación Si R es un anillo, entonces el conjunto de todos los polinomios con coeficientes en R es denotado por $R[x]$.

Proposición 2.4. *Si R es un anillo, entonces $R[x]$ es un anillo que contiene a R como un subanillo.*

Demostración. Si $\sigma = (s_0, s_1, ...)$, $\tau = (t_0, t_1, ...)$ y $\gamma = (v_0, v_1, ...)$ definimos la suma y la multiplicación respectivamente como,

$$\sigma+\tau=(s_0 + t_0, s_1 + t_1, ..., s_n + t_n, ...)$$

y

$$\sigma\tau= (c_0, c_1, c_2, ...),$$

donde c_k esta dada por

$$c_k= \sum_{i=0}^{k} s_i t_{k-i}= \sum_{i+j=k} s_i t_j$$

Veamos que $R[x]$ es un anillo.

1. Inverso aditivo.
   Si $\sigma = (s_0, s_1, ..., s_n, ...) \in R[x]$, definimos $-\sigma = (-s_0, -s_1, ..., -s_n, ...) \in R[x]$ tal que $\sigma+(-\sigma)= (s_0+(-s_0), s_1+(-s_1), ..., s_n+(-s_n), ...)= (0, 0, ...,0, ...)$.

2. Neutro aditivo.
   Si $\sigma = (s_0, s_1, ..., s_n, ...) \in R[x]$, y $\mathbf{0} = (0, 0, 0, ...) \in R[x]$, entonces $\sigma + \mathbf{0} = (s_0 + 0, s_1 + 0, ..., s_n + 0, ...) = (s_0, s_1, ..., s_n, ...)$.

3. Conmutatividad de la suma.
   Sean $\sigma, \tau \in R[x]$, entonces $\sigma + \tau = (s_0, s_1, ..., s_n, ...) + (t_0, t_1, ..., t_n, ...) = (s_0+t_0, s_1+t_1, ..., s_n+t_n, ...) = (t_0+s_0, t_1+s_1, ..., t_n+s_n, ...) = (t_0, t_1, ..., t_n, ...)+ (s_0, s_1, ..., s_n, ...) = \tau + \sigma$.

4. Asociatividad.
   Sean $\sigma, \tau, \gamma \in R[x]$ entonces $\sigma + (\tau + \gamma) = (q_0, q_1, ..., q_n)$ donde $q_i = s_i + (t_i + v_i) = (s_i + t_i) + v_i$ para cada i. Por lo tanto $\sigma + (\tau + \gamma) = (\sigma + \tau) + \gamma$.

5. Conmutatividad del producto.
   Sea $\sigma, \tau \in R[x]$, luego $\sigma\tau = (c_0, c_1, ...)$ donde

$$c_k = \sum\nolimits_{i+j=k} s_i t_j = \sum\nolimits_{j+i=k} t_j s_i$$

   por lo tanto $\sigma\tau = \tau\sigma$.

6. Asociatividad del producto.
   Sea $\sigma, \tau, \gamma \in R[x]$ con $\gamma = (r_0, r_1, ..., r_n)$ entonces $\gamma(\sigma\tau) = \gamma(c_0, c_1, ..., c_n, ...)$ donde

$$c_k = \sum\nolimits_{i+j=k} s_i t_j$$

   luego $\gamma(c_0, c_1, ..., c_n, ...) = (d_0, d_1, ..., d_n, ...)$ donde

$$d_l = \sum\nolimits_{n+k=l} r_n c_k = \sum r_n (\sum\nolimits_{i+j=k} s_i t_j) = \sum\nolimits_{h+i+j=l} r_n (s_i t_j) =$$

$$\sum\nolimits_{i+h+j=l} (r_n s_i) t_j = \sum\nolimits_{k'+j} c'_k t_j$$

   donde

$$c'_k = \sum (r_n s_i).$$

   Por lo tanto $\gamma(\sigma\tau) = (\gamma\sigma)\tau$.

7. Distributividad del producto.
   Sea $\sigma, \tau, \gamma \in R[x]$. Entonces $\sigma(\tau + \gamma) = (c_0, c_1, ..., c_k, ...)$ donde

$$c_k = \sum\nolimits_{i+j=k} s_i (t_j + r_j) = \sum\nolimits_{i+j=k} s_i t_j + \sum\nolimits_{i+j=k} s_i r_j$$

   donde es la entrada k-ésima de $\sigma\tau + \sigma\gamma$. Por lo tanto $\sigma(\tau + \gamma) = \sigma\tau + \sigma\gamma$.

8. Neutro respecto al producto.
   Afirmamos que $1_{R[x]} = (1, 0, 0, ..., 0, ...)$ tal que si $\sigma = (s_0, s_1, ..., s_n, ...) \in R[x]$ entonces $\sigma 1_{R[x]} = (c_0, c_1, ..., c_k, ...)$ si identificamos $1_{R[x]} = (1, 0, 0, ..., 0, ...) = (t_0, t_1, ...t_n, ...)$ se sigue que

$$c_k = \sum\nolimits_{i+j=k} s_i t_j = s_k t_0 = s_k$$

   Por lo tanto $\sigma 1_{R[x]} = (s_0, s_1, ..., s_n, ...)$

Observemos que el subconjunto $\{(r, 0, 0, ...) \mid r \in R\}$ es un subanillo de $R[x]$ que identificamos con R.

□

Lema 2.5. *Sea R un anillo y sea $\sigma, \tau \in R[x]$ polinomios distintos de cero,*

I. $\sigma\tau = \boldsymbol{0}$ *ó* $\boldsymbol{gr}(\sigma\tau) \leq \boldsymbol{gr}(\sigma) + \boldsymbol{gr}(\tau)$.

II. *Si R es un dominio entero, entonces* $\sigma\tau \neq \boldsymbol{0}$ *y* $\boldsymbol{gr}(\sigma\tau) = \boldsymbol{gr}(\sigma) + \boldsymbol{gr}(\tau)$.

III. *Si R es un dominio entero, entonces $R[x]$ es un dominio entero.*

Demostración. I. Sea $\sigma = (s_0, s_1, ..., s_n, ...)$, $\tau = (t_0, t_1, ..., t_m, ...) \in R[x]$ con $\mathbf{gr}(\sigma) = n$ y $\mathbf{gr}(\tau) = m$ respectivamente. Sea $\sigma\tau \neq \mathbf{0}$, y supongamos que $\mathbf{gr}(\sigma\tau) = k > n + m = \mathbf{gr}(\sigma) + \mathbf{gr}(\tau)$. Consideremos

$$c_k = \sum_{i=1}^{k} s_i t_{k-i}$$

donde c_k es la entrada k-ésima de $\sigma\tau$, y así tenemos dos casos:

1. Si $i > n$ entonces $s_i = 0$, por lo que $s_i t_{k-i} = 0$.
2. Observemos que si $i \leq n$ implica $0 \leq n - i$, luego considerando que $k > n + m$, tenemos que $k - i > n + m - i \geq m$. Por lo que $t_{k-i} = 0$ y así $s_i t_{k-i} = 0$.

Esto implica que $c_k = 0$ para todo $k > n + m$, contradicción. Por lo tanto $k \leq n + m$.

II. Afirmamos que cada término en

$$c_{n+m} = \sum_{i=1}^{n+m} s_i t_{n+m-i}$$

es 0 ya que:

1. Si $i > n$, entonces $s_i = 0$. Por lo tanto $s_i t_{n+m-i} = 0$.
2. Si $i < n$, entonces $0 < n - i$, luego $m < n + m - i$, por lo que $t_{n+m-i} = 0$. Por lo tanto $s_i t_{n+m-i} = 0$.

con la excepción de $i = n$, que implica $c_{n+m} = s_n t_m$, ya que R es dominio entero, $s_n \neq 0$ y $t_n \neq 0$ implica que $c_{n+m} \neq 0$, donde c_{n+m} es el coeficiente principal de $\sigma\tau$, de ahí que $\sigma\tau \neq 0$, además $\mathbf{gr}(\sigma\tau) = \mathbf{gr}(\sigma) + \mathbf{gr}(\tau)$.

III. Es consecuencia del inciso 2. Puesto que el producto de dos polinomios distintos de cero, es distinto de cero.

□

Definición 2.6. *Si R es un anillo, entonces $R[x]$ es llamado el* ***anillo de polinomios*** *sobre R.*

Definición 2.7. *Definimos el elemento $x \in R[x]$ por*

$$x = (0, 1, 0, 0, ...).$$

Proposición 2.8. *1. Si $\tau = (t_0, t_1, ..., t_n, ...) \in R[x]$, entonces*

$$x\tau = (0, t_0, t_1, ...).$$

2. Si $n \geq 1$, entonces x^n es el polinomio que tiene 0 en todos los lugares excepto el 1 en el n-ésima coordenada.

3. Si $r \in R$, entonces

$$(r, 0, 0, ...)(s_0, s_1, ..., s_j, ...) = (rs_0, rs_1, ..., rs_j, ...)$$

Demostración. 1. Sean $x = (0, 1, 0, 0, ...), \tau = (t_0, t_1, ..., t_n, ...) \in R[x]$, luego $x\tau = (c_0, ..., c_k, ...)$. Si identificamos $x = (0, 1, 0, 0, ...) = (s_0, s_1, ...)$, entonces $c_0 = \sum_{i+j=0} s_i t_j = s_0 t_0 = 0$, y además para $k > 0$. $c_k = \sum_{i+j=k} s_i t_j = s_1 t_{k-1} = t_{k-1}$. Por lo tanto $x\tau = (c_0, c_1, ..., c_k, ...) = (0, t_0, t_1, ..., t_{k-1}, ...)$.

2. Por inducción, para $n = 1$ entonces $x = (0, 1, 0, 0, ...)$, se cumple. Supongamos que se cumple para $n = k$, es decir $x^k = (0, ..., 0, 1, 0, ...)$, donde hay k ceros a la izquierda del 1. Veamos que se cumple para $x^{k+1} = x^k x^1 = (0, ..., 0, 1, 0, ...)(0, 1, 0, 0, ...) = (0, ..., 0, 0, 1, 0, ...)$ donde hay $k + 1$ ceros a la izquierda.

3. Sean $r = (r, 0, 0, 0, ...), \tau = (t_0, t_1, ..., t_n, ...) \in R[x]$. Si identificamos a $r = (r, 0, 0, 0, ...) = (s_0, s_1, ...)$, entonces $r\sigma = (c_0, c_1, ..., c_k, ...)$, donde $c_k = \sum_{i=0}^{k} s_i t_{k-i} = s_0 t_k = r t_k$. Por lo tanto $r\sigma = (c_0, ..., c_k, ...) = (rt_0, ..., rt_n, ...)$.

□

Proposición 2.9. *Si* $\sigma = (s_0, s_1, ..., s_n, 0, 0, ...)$*, entonces*

$$\sigma = s_0 + s_1x + s_2x^2 + ... + s_nx^n.$$

donde cada término $s \in R$ *esta identificado con el polinomio* $(s, 0, 0, ...)$

Demostración. $\sigma = (s_0, s_1, ..., s_n, 0, 0, ...)$

$= (s_0, 0, 0, ...) + (0, s_1, 0, ...) + (0, 0, 0, ..., s_n, 0...)$

$= s_0(1, 0, 0, ...) + s_1(0, 1, 0, ...) + s_n(0, 0, 0, ..., 1, 0...)$

$= s_0 + s_1x + s_2x^2 + ... + s_nx^n$

□

Notación: Identificamos a σ con $f(x) = s_0+s_1x+s_2x^2+...+s_nx^n$ donde s_0 el término constante, s_n el coeficiente principal. Si el coeficiente principal $s_n = 1$, entonces $f(x)$ es llamado **mónico** y $n = \mathbf{gr}(f)$. Un polinomio constante es el polinomio cero ó un polinomio de grado 0.

Definición 2.10. *Si* R *es un anillo y* $f(x) = \sum_{i=0}^{n} s_ix^i \in R[x]$ *con* $\boldsymbol{gr}(f(x)) = n \geq 1$*, definimos su derivada* $f'(x) \in R[x]$ *por*

$$f'(x) = s_1 + 2s_2x + 3s_3x^2 + ... + ns_nx^{n-1};$$

si $f(x)$ *es un polinomio constante, definimos su derivada como el polinomio cero.*

Además se cumple lo siguiente:

1. $(f + g)' = f' + g'$;
2. $(rf)' = r(f')$ si $r \in R$;
3. $(fg)' = fg' + f'g$;
4. $(f^n)' = nf^{n-1}f'$ para todo $n \geq 1$.

Proposición 2.11. *Sea* R *un anillo y sea* $f(x) \in R[x]$*.*

1. *Si* $(x - a)^2 \mid f(x)$*, entones* $x - a \mid f'(x)$ *en* $R[x]$*.*
2. *Si* $x - a \mid f(x)$ *y* $x - a \mid f'(x)$*, entonces* $(x - a)^2 \mid f(x)$*.*

Demostración. 1. Sea $f(x) \in R[x]$ y supongamos $(x-a)^2 \mid f(x)$ entonces, existe $g(x) \in R[x]$ tal que $f(x) = g(x)(x-a)^2$. Luego,

$$\begin{aligned} f'(x) &= g'(x)(x-a)^2 + g(x)(2)(x-a) \\ &= (x-a)[g'(x)(x-a) + 2g(x)]. \end{aligned}$$

Así $x - a \mid f'(x)$.

2. Sean $x-a \mid f(x)$ y $x-a \mid f'(x)$, entonces existen $g(x), h(x) \in R[x]$ tales que $f(x) = g(x)(x-a)$ y $f'(x) = h(x)(x-a)$, de la primera igualdad $f'(x) = g'(x)(x-a) + g(x)$. Entonces

$$h(x)(x-a) = g'(x)(x-a) + g(x)$$

luego,

$$h(x)(x-a) - g'(x)(x-a) = g(x).$$

Así

$$f(x) = (x-a)[h(x) - g'(x)](x-a).$$

De ahí que $f(x) = (x-a)^2[h(x) - g'(x)]$, por lo tanto $(x-a)^2 \mid f(x)$. □

2.1. Máximo Común Divisor

Proposición 2.12. *Sea $\mathbb{F}$ un campo y $f(x), g(x) \in \mathbb{F}[x]$ donde $f(x) \neq 0$. Entonces existen únicos $q(x), r(x) \in \mathbb{F}[x]$ tales que $g(x) = q(x)f(x) + r(x)$ con $r(x) = \mathbf{0}$ ó $\mathbf{gr}(r(x)) < \mathbf{gr}(f(x))$.*

Demostración. **Existencia** . Si $f(x) \mid g(x)$ en $\mathbb{F}[x]$, entonces $g(x) = q(x)f(x)$ para algún $q(x) \in \mathbb{F}[x]$ y con $r(x) = \mathbf{0}$. Si $f(x) \nmid g(x)$. Sea $\mathcal{A} = \{g(x) - q(x)f(x) \neq 0 \mid q(x) \in \mathbb{F}[x]\}$. Notemos que $\mathcal{A} \neq \emptyset$ ya que si $\mathcal{A} = \emptyset$ implica que $f(x) \mid g(x)$. Luego, por el Principio del Buen Orden existe $r(x) \in \mathcal{A}$ de menor grado tal que $r(x) = g(x) - q(x)f(x)$ para algún $q(x) \in \mathbb{F}[x]$, entonces $r(x) = g(x) - q(x)f(x) \neq 0$. Solo resta demostrar que $\mathbf{gr}(r(x)) < \mathbf{gr}(f(x))$. Supongamos que $\mathbf{gr}(f(x)) \leq \mathbf{gr}(r(x))$. Sean $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_0$

y $r(x) = b_m x^m + b_{m-1}x^{m-1} + ... + b_0$ con $\mathbf{gr}(f(x)) = n$ y $\mathbf{gr}(r(x)) = m$ respectivamente. Si $n \leq m$ se sigue que $x^{m-n} \in \mathbb{F}[x]$. Definimos $h(x) = r(x) - a_n^{-1}b_m x^{m-n} f(x) \in \mathbb{F}[x]$. Notemos que $h(x) = 0$ ó $\mathbf{gr}(h(x)) < \mathbf{gr}(r(x))$. Supongamos que $h(x) = 0$. Entonces $r(x) = a_n^{-1}b_m x^{m-n} f(x)$, además $g(x) = q(x)f(x) + r(x) = (q(x) + a_n^{-1}b_m x^{m-n})f(x)$ lo que contradice $f(x) \nmid g(x)$. Si $h(x) \neq 0$, entonces $\mathbf{gr}(h(x)) < \mathbf{gr}(r(x))$ y $g(x) - q(x)f(x) = r(x) = h(x) + a_n^{-1}b_m x^{m-n} f(x)$ se sigue que $h(x) = g(x) - (q(x) + a_n^{-1}b_m x^{m-n})f(x) \in \mathcal{A}$ contradiciendo el hecho de que $r(x)$ es un polinomio de menor grado en $\mathcal{A}$. Por lo tanto $\mathbf{gr}(r(x)) < \mathbf{gr}(f(x))$.
Unicidad. Supongamos que exiten $q'(x), r'(x) \in \mathbb{F}[x]$ tal que $g(x) = q'(x)f(x) + r'(x)$, con $r'(x) = 0$ ó $\mathbf{gr}(r'(x)) < \mathbf{gr}(f(x))$, entonces $q(x)f(x) + r(x) = q'(x)f(x) + r'(x)$, se sigue que $(q(x) - q'(x))f(x) = r'(x) - r(x)$. Suponga-mos que $r'(x) \neq r(x)$ entonces cada lado tiene un grado. Pero el $\mathbf{gr}((q(x) - q'(x))f(x)) = \mathbf{gr}(q(x) - q'(x)) + \mathbf{gr}(f(x)) \geq \mathbf{gr}(f(x))$, mientras que $\mathbf{gr}(r'(x) - r(x)) \leq max\{\mathbf{gr}(r'(x)), \mathbf{gr}(r(x))\} < \mathbf{gr}(f(x))$, contradicción. Por lo tanto $r'(x) = r(x)$ y $(q(x) - q'(x))f(x) = 0$. Como $\mathbb{F}[x]$ es un dominio entero y $f(x) \neq 0$, se sigue que $q(x) - q'(x) = 0$ por lo tanto $q(x) = q'(x)$.

□

Definición 2.13. *Sea $f(x) = a_0 + a_1 x + ... + a_n x^n \in R[x]$. Un polinomio define una función polinomial $f : R \longrightarrow R$, tal que $a \mapsto f(a) = a_0 + a_1 a + a_2 a^2 + ... + a_n a^n$*

Definición 2.14. *Sea $\mathbb{F}$ un campo y $f(x) \in \mathbb{F}[x]$. Un elemento $a \in \mathbb{F}$ es una raíz de $f(x)$ si $f(a) = 0$.*

Lema 2.15. *Sea $\mathbb{F}$ un campo, $f(x) \in \mathbb{F}[x]$ y $a \in \mathbb{F}$. Entonces existe un polinomio $q(x) \in \mathbb{F}[x]$ tal que $f(x) = q(x)(x - a) + f(a)$.*

Demostración. Por la Proposición 2.12, existen $q(x), r(x) \in \mathbb{F}[x]$, tales que $f(x) = q(x)(x - a) + r(x)$ con $r(x) = 0$ ó $\mathbf{gr}(r(x)) < \mathbf{gr}(x - a) = 1$. Por lo tanto r es una constante porque $x - a$ tiene grado 1. Luego evaluando:

$$f(a) = q(a)(a - a) + r.$$

Así $r = f(a)$. Por lo tanto $f(x) = q(x)(x - a) + f(a)$. □

Proposición 2.16. *Sea $\mathbb{F}$ un campo, $f(x) \in \mathbb{F}[x]$ y $a \in \mathbb{F}$. Entonces $a \in \mathbb{F}$ es raíz de $f(x)$ si y sólo si $(x - a) \mid f(x)$ en $\mathbb{F}[x]$.*

Demostración. $\Rightarrow$] Por el Lema 2.15, $f(x) = q(x)(x-a) + f(a)$ para algún $q(x) \in \mathbb{F}[x]$, además $f(a) = 0$. Por lo tanto $f(x) = q(x)(x-a)$, es decir, $(x-a) \mid f(x)$.
$\Leftarrow$] Por hipótesis, existe $h(x) \in \mathbb{F}[x]$ tal que $f(x) = h(x)(x-a)$, evaluando en a, $f(a) = h(a)(a-a) = 0$. Por lo tanto a es una raíz de $f(x)$ en $\mathbb{F}$. □

Proposición 2.17. *Sea $\mathbb{F}$ un campo y $0 \neq f(x) \in \mathbb{F}[x]$ con $\boldsymbol{gr}(f(x)) = n$. Entonces $f(x)$ tiene a lo más n raíces en $\mathbb{F}$.*

Demostración. La demostración se hará por inducción sobre $n = \mathbf{gr}(f(x))$. Para $n = 0$, tenemos $0 \neq f(x) = c \in \mathbb{F}$ con c una constante, entonces el número de raíces de $f(x)$ es cero que es menor o igual al $\mathbf{gr}(f) = 0$. Ahora sea $n \geq 1$. Si $f(x)$ no tiene raíces en $\mathbb{F}$, entonces el número de raíces de $f(x)$ es cero, que es menor al $\mathbf{gr}(f(x))$. Supongamos que $f(x)$ tiene a $a \in \mathbb{F}$ como raíz. Entonces por la Proposición 2.16, existe $q(x) \in \mathbb{F}[x]$ tal que $f(x) = q(x)(x-a)$; más aún, $q(x)$ tiene grado $n-1$. Ahora sea $a \neq b \in \mathbb{F}$ una raíz de $f(x)$, entonces

$$0 = f(b) = q(b)(b-a).$$

Ya que $b \neq a$, entonces $q(b) = 0$, por lo tanto b es una raíz de q. Como el $\mathbf{gr}(q(x)) = n - 1 < n$, así que por la hipótesis de inducción que afirma que $q(x)$ tiene a lo más $n-1$ raíces en $\mathbb{F}$. Se concluye que $f(x)$ tiene a lo más n raíces en $\mathbb{F}$. □

Definición 2.18. *Sea $\mathbb{F}$ un campo y $f(x), g(x) \in \mathbb{F}[x]$. Un polinomio $c(x) \in \mathbb{F}[x]$ es un divisor común de $f(x)$ y $g(x)$, si $c(x) \mid f(x)$ y $c(x) \mid g(x)$ en $\mathbb{F}[x]$. Si $f(x)$ y $g(x)$ no ambos cero, escribimos $m.c.d(f(x), g(x))$ para denotar el máximo común divisor de $f(x)$ y $g(x)$, y lo definimos como un divisor común de $f(x)$ y $g(x)$ tal que es un polinomio mónico de mayor grado.*

Si $f(x) = g(x) = 0$, entonces $m.c.d(f(x), g(x)) = 0$.

Proposición 2.19. *Sea $\mathbb{F}$ un campo, $f(x), g(x) \in \mathbb{F}[x]$. Entonces $d(x) = m.c.d(f(x), g(x))$ existe y es combinación lineal de $f(x)$ y $g(x)$. Es decir, existen $s(x), t(x) \in \mathbb{F}[x]$ tales que $d(x) = s(x)f(x) + t(x)g(x)$.*

Una vez introducido el concepto de dominio de ideales principales podremos demostrar esta proposición.

Definición 2.20. *Sean R un dominio entero y $p \in R$. Decimos que p es un elemento irreducible en R si $p \neq 0$, p no es unidad y $p = ab$ con $a, b \in R$, implica que $a \in R$ es unidad o $b \in R$ es unidad.*

Lema 2.21. *Un polinomio* $0 \neq g(x) \in \mathbb{F}[x]$ *es unidad si y sólo si* $\boldsymbol{gr}(g(x)) = 0$

Demostración. $\Rightarrow$] Si $g(x)$ es unidad, entonces existe $h(x) \in \mathbb{F}[x]$ tal que $g(x)h(x) = 1$, de modo que $\mathbf{gr}(g(x))+\mathbf{gr}(h(x)) = 0$ implica $\mathbf{gr}(g(x)) = 0$, por lo tanto $g(x) \in \mathbb{F}$.
$\Leftarrow$] Si $\mathbf{gr}(g(x)) = 0$, entonces $0 \neq g(x) \in \mathbb{F}$. Por lo tanto $g(x)$ es unidad. □

Proposición 2.22. *Si* $\mathbb{F}$ *es un campo. Entonces un elemento* $p(x) \in \mathbb{F}[x]$ *es irreducible si y sólo si* $\boldsymbol{gr}(p(x)) = n \geq 1$ *y no existe una factorización en* $\mathbb{F}[x]$ *de la forma* $p(x) = g(x)h(x)$ *con* $0 < \boldsymbol{gr}(g(x)) < n$ *y* $0 < \boldsymbol{gr}(h(x)) < n$.

Demostración. $\Rightarrow$] Sea $p(x) \in \mathbb{F}[x]$ irreducible, así $p(x) \neq 0$ y $p(x)$ no es unidad en $\mathbb{F}[x]$, implica $\mathbf{gr}(p(x)) \geq 1$. Si $p(x) = g(x)h(x)$, entonces $g(x)$ es unidad y así $\mathbf{gr}(g(x)) = 0$ y $\mathbf{gr}(h(x)) = n$ ó $h(x)$ es unidad implica $\mathbf{gr}(g(x)) = n$ y $\mathbf{gr}(h(x)) = 0$.
$\Leftarrow$] Notemos que $\mathbf{gr}(p(x)) = n \geq 1$ implica que $p(x) \neq 0$ y $p(x)$ no es unidad en $\mathbb{F}[x]$. Sea $p(x) = a(x)b(x)$ en $\mathbb{F}[x]$ y supongamos que $a(x)$ no es unidad y que $b(x)$ no es unidad, entonces $\mathbf{gr}(a(x)) > 0$ y $\mathbf{gr}(b(x)) > 0$ además $n = \mathbf{gr}(p(x)) = \mathbf{gr}(a(x)) + \mathbf{gr}(b(x))$ implica $\mathbf{gr}(a(x)) < n$ y $\mathbf{gr}(b(x)) < n$, contradicción. Por lo tanto $a(x) \in \mathbb{F}[x]$ es unidad o $b(x) \in \mathbb{F}[x]$. □

Lema 2.23. *Sea* $\mathbb{F}$ *un campo y* $p(x), f(x) \in \mathbb{F}[x]$ *con* $p(x)$ *irreducible mónico. Entonces*

$$d(x) = m.c.d.(p(x), f(x)) = \begin{cases} 1 & , p(x) \nmid f(x) \\ p(x) & , p(x) \mid f(x) \end{cases}$$

Demostración. Como $d(x) \mid p(x)$, entonces existe $q(x) \in \mathbb{F}[x]$, tal que $p(x) = d(x)q(x)$, así $d(x)$ es unidad o $q(x)$ es unidad. Si $d(x)$ es unidad, entonces $\mathbf{gr}(d(x)) = 0$, luego $d(x) = c \in \mathbb{F}[x]$ un polinomio constante, esto significa que c es el coeficiente principal de $d(x)$. Por lo tanto $d(x) = 1$ por ser mónico. Si $q(x)$ es unidad, $q(x) = u \in K$, así $p(x) = d(x)u$, el coeficiente principal de $p(x)$ es 1, el coeficiente principal de $d(x)u = u$ ya que $d(x)$ es mónico. Por lo tanto $u = 1$ y así $d(x) = p(x)$. Por lo tanto

$$d(x) = \begin{cases} 1 \\ p(x) \end{cases}$$

Luego $p(x) \mid f(x)$ y $p(x) \mid p(x)$, entonces $p(x) \mid d(x)$ y $d(x) \mid p(x)$. Así $d(x) = p(x)u(x)$, $p(x) = d(x)v(x)$ con $u(x), v(x) \in \mathbb{F}[x]$, así $d(x) = d(x)v(x)u(x)$

implica que $u(x)v(x) = 1$. Por lo tanto $u(x) \in \mathbb{F}[x]$ es unidad. Como $d(x) = p(x)u(x)$ tenemos que $d(x) = p(x)$ ya que $p(x), d(x)$ son mónicos.
Si $p(x) \nmid f(x)$, podemos suponer que $d(x) = p(x)$ así $d(x) \mid f(x)$ por lo tanto $p(x) \mid f(x)$, lo que es una contradicción. Por lo tanto $d(x) = 1$. □

Definición 2.24. *Dos polinomios $f(x), g(x) \in \mathbb{F}[x]$, donde $\mathbb{F}$ es un campo, son llamados **primos relativos** si $m.c.d(f(x), g(x)) = 1$.*

Proposición 2.25. *Sea $f(x) = (x - a_1)...(x - a_n) \in \mathbb{F}[x]$ con $\mathbb{F}$ un campo, entonces $f(x)$ no tiene raíces repetidas si y sólo si $m.c.d(f, f') = 1$.*

Demostración. $\Rightarrow$] Supongamos que $f(x)$ no tiene raíces repetidas y supongamos que $m.c.d(f, f') = d(x)$ con $d(x) \neq 1$. Como $\mathbb{F}[x]$ es campo, entonces $\mathbf{gr}(d(x)) > 0$. Sea $(x - a)$ un factor de $d(x)$, por la Proposición 2.11 parte 2. $x - a \mid f'(x)$ entonces $(x - a)^2 \mid f(x)$. Luego a tiene multiplicidad 2, entonces a es una raíz repetida. Por lo tanto $f(x)$ tiene raíces repetidas, contradicción. Así $d(x) = 1$.
$\Leftarrow$] Supongamos que $f(x)$ tiene una raíz repetida, digamos a, entones $(x-a)^2$ por la Proposición 2.11 parte 1. $(x-a) \mid f'(x)$, entonces $(x-a) \mid m.c.d(f, f)'$ así $(x-a) \mid 1$. Por lo que $(x-a)$ es unidad, contradicción, ya que $\mathbf{gr}(x-a) = 1$. Por lo tanto $f(x)$ no tiene raíces repetidas. □

Proposición 2.26. *Sea $\mathbb{F}$ y $f(x) \in \mathbb{F}[x]$ tal que $\boldsymbol{gr}(f(x)) \geq 1$. Entonces existe $0 \neq a \in \mathbb{F}$ y $p_1(x), ..., p_r(x) \in \mathbb{F}[x]$ mónicos irreducibles tales que $f(x) = ap_1(x)...p_2(x)$, luego esta factorización es única salvo por el orden de los factores.*

Demostración. **Existencia.** Por inducción en el $\mathbf{gr}(f(x)) = n \geq 1$. Si $n = 1$ entonces $f(x) = ax+b= a(x+ba^{-1})$ y con $a \neq 0$, luego $x+ba^{-1} \in \mathbb{F}[x]$ es mónico irreducible con $a \in \mathbb{F}$. Sea $n > 1$, si $f(x) \in \mathbb{F}$ es irreducible, con a es el coeficiente principal de $f(x)$, entonces $f(x) = a(a^{-1}f(x))$ donde $a^{-1}f(x) \in \mathbb{F}[x]$ es irreducible mónico, si $f(x) \in \mathbb{F}[x]$ no es irreducible existen $g(x), h(x) \in \mathbb{F}[x]$ tales que $f(x) = g(x)h(x)$ y $0 < \mathbf{gr}(h(x)), \mathbf{gr}(g(x)) < n$. Por hipótesis inductiva existen $a, b \in \mathbb{F}$ no cero y $p_1(x), ..., p_l(x), q_1(x), ..., q_s(x) \in \mathbb{F}[x]$ polinomios mónicos irreducibles tal que $g(x) = ap_1(x), ..., p_l(x)$ y $h(x) = bq_1(x), ..., q_s(x)$, por lo tanto $f(x) = abp_1(x)...p_l(x)q_1(x)...q_s(x)$ con $0 \neq ab \in \mathbb{F}$.
Unicidad. Sea $f(x) = ap_1(x)...p_l(x)=bq_1(x)...q_r(x)$ con $a, b \in \mathbb{F}$ no ceros y cada $p_i, q_i \in \mathbb{F}[x]$ irreducibles mónicos, por lo tanto $a = b$. Luego $p_1(x)...p_l(x) = q_1(x)...q_r(x)$, sea $M = max\{l, r\} \geq 1$, se hará inducción sobre $M \geq 1$, si $M = 1$, entonces $ap(x) = bq_1$, y por lo tanto $p_1(x) = q_1(x)$. Si

$M > 1$ entonces $p_l \mid p_1(x), ..., p_l(x) = q_1(x), ..., q_r(x)$, así existe $j \in \{1, ..., r\}$ tal que $p_l(x) \mid q_j(x)$ entonces $q_j = h(x)p_l(x)$ lo cual implica $h(x) \in \mathbb{F}$, sin perdida de generalidad $q_j(x) = q_r(x) = p_l(x)$, asi $h(x) = 1$ y $q_r(x) = p_l(x)$, por lo que $p_1(x), ..., p_{l-1}(x) = q_1(x), ..., q_{r-1}(x)$. Por la hipótesis de inducción $l - 1 = r - 1$ si y sólo si $l = r$, por lo tanto $p_i(x) = q_i(x)$ para todo $i \in \{1, ..., l - 1\}$. □

Capítulo 3

Morfismos

En el estudio de cualquier teoría matemática formal hay dos conceptos muy importantes, uno es el estudio de los objetos y el otro el estudio de las relaciones entre los objetos. En la teoría de los anillos conmutativos finitos los objetos son los anillos y las relaciones entre ellos más importantes son los morfismos. En esta sección daremos cuenta de las propiedades más importantes de estos morfismos necesarias para desarrollar la teoría.

Definición 3.1. *Sean R y R' anillos. Un morfismo de anillos es una función $f : R \to R'$ que cumple lo siguiente :*

- $f(r_1 + r_2) = f(r_1) + f(r_2)$
- $f(r_1 \cdot r_2) = f(r_1) \cdot f(r_2)$
- $f(1_R) = 1_{R'}$

Lema 3.2. $f(0) = 0_{R'}$

Demostración. Notemos que:
$0_{R'} + f(0) = f(0) = f(0 + 0) = f(0) + f(0)$.
Cancelando en ambos lados $f(0)$, se obtiene el resultado. □

Definición 3.3. *Un morfismo de anillos $f : R \to R'$ es llamado **monomorfismo** si para cualesquiera $g_1, g_2 : R'' \to R$ morfismos de anillos tal que $fg_1 = fg_2$, implica que $g_1 = g_2$.*

Definición 3.4. *Un morfismo de anillos $f : R \to R'$ es llamado **epimorfismo** si para cualesquiera $g_1, g_2 : R' \to R''$ morfismos de anillos tal que $g_1 f = g_2 f$, implica que $g_1 = g_2$.*

Proposición 3.5. *Sean R y R' anillos y $\gamma : R \to R'$ un morfismo. Entonces γ es monomorfismo si y sólo si γ es inyectiva.*

Demostración. $\Rightarrow$] Supongamos que $f(a) = f(b)$, para $a, b \in R$ y $a \neq b$. Consideremos el anillo de polinomios con coeficientes en $\mathbb{Z}$, $\mathbb{Z}[x]$. Y tomemos los morfismos $h, g : \mathbb{Z}[x] \to R$ tal que $h(x) = a$ y $g(x) = b$. Así

$$\begin{aligned} fh(z_0 + z_1x + ... + z_nx^n) &= f(h(z_0 + z_1x + ... + z_nx^n)) \\ &= f(z_0 + z_1a + ... + z_na^n) \\ &= z_0 + z_1f(a) + ... + z_nf(a)^n \\ &= z_0 + z_1f(b) + ... + z_nf(b)^n \\ &= f(z_0 + z_1b + ... + z_nb^n) \\ &= fg(z_0 + z_1x + ... + z_nx^n). \end{aligned}$$

Por lo tanto $fh = fg$, entonces $h = g$, es decir, $a = h(x) = g(x) = b$.
$\Leftarrow$] Sea $f = R \to R'$ y $g, h : R'' \to R$ morfismos, tales que $fg = fh$. Por demostrar que $g = h$.
Sea $r \in R''$ tal que $(fg)(r) = (fh)(r)$, entonces $f(g(r)) = f(h(r))$. Por lo tanto $g(r) = h(r)$ para todo $r \in R''$. □

Proposición 3.6. *Sea $f : A \to R$ un morfismo de anillos, sea $0 \neq a \in A$, entonces*

1. *$f(a^n) = (f(a))^n$ para todo $n \in \mathbb{N}$.*
2. *Si $a \in U(A)$, entonces $f(a^{-1}) = (f(a))^{-1}$. Además $f(a^{-n}) = (f(a))^{-n}$ para todo $n \in \mathbb{N}$.*

Demostración.

1. Por inducción sobre n. Si $n = 0$, entonces $a^n = 1$, luego $f(1) = 1$, además $(f(a))^n = 1$. Por lo tanto $f(a^n) = (f(a))^n$ para $n = 0$. Para el paso inductivo. $f(a^{n+1}) = f(aa^n) = f(a)f(a^n) = f(a)(f(a))^n = (f(a))^{n+1}$.

2. Si $a \in U(a)$, entonces existe $a^{-1} \in A$, tal que $aa^{-1} = 1$, entonces $f(a)f(a^{-1}) = f(aa^{-1}) = f(1) = 1$. Entonces $f(a^{-1}) = (f(a))^{-1}$.
   Notemos que $f(U(A)) \subseteq U(R)$.
   Si $n > 0$, $a^{-n} = (a^{-1})^n$. Luego $f(a^{-n}) = f((a^{-1})^n) = f(a^{-1})^n = (f(a)^{-1})^n = (f(a))^{-n}$.

□

Lema 3.7. *El morfismo $i : \mathbb{Z} \hookrightarrow \mathbb{Q}$ es epimorfismo pero no es suprayectivo.*

Demostración. Sean $\alpha, \beta : \mathbb{Q} \to R$ morfismos de anillos tales que $\alpha i = \beta i$. Sea $m \in \mathbb{Z}$, entonces $\alpha(m) = \alpha i(m) = \beta i(m) = \beta(m)$. Es decir, $\alpha(m) = \beta(m)$, para todo $m \in \mathbb{Z}$.
Sea $0 \neq q \in \mathbb{Q}$ tal que $q = nm^{-1}$ para $n, m \in \mathbb{Z}$ y $m \neq 0$, entonces

$$\begin{aligned}\alpha(q) = \alpha(nm^{-1}) &= \alpha(n)\alpha(m^{-1}) \\ &= \alpha(n)\alpha(m)^{-1} \\ &= \beta(n)\beta(m)^{-1} \\ &= \beta(n)\beta(m^{-1}) \\ &= \beta(nm^{-1}) \\ &= \beta(q).\end{aligned}$$

Por lo tanto $\alpha(q) = \beta(q)$ para todo $q \in \mathbb{Q}$. Por lo tanto $\alpha = \beta$. Así que i es epimorfismo y claramente no es suprayectiva. Por lo que en teoría de anillos epimorfismo no coincide con morfismo suprayectivo. □

Definición 3.8. *Sea $f : R \to R'$ un morfismo, f es un **isomorfismo** si existe un morfismo $g : R' \to R$, tal que $f \circ g = 1_{R'}$ y $g \circ f = 1_R$.*

Proposición 3.9. *Son equivalentes los siguientes enunciados:*

1. *$f : R \to S$ es isomorfismo de anillos.*
2. *$f : R \to S$ es morfismo biyectivo.*

Demostración. 1. $\Rightarrow$] 2. Como $f : R \to S$ es isomorfismo en particular es morfismo, además existe $g : S \to R$ tal que $f \circ g = 1_S$, $g \circ f = 1_R$ entonces f es suprayectiva y f es inyectiva, por lo tanto f es biyectica.
2. $\Rightarrow$] 1. Como f es biyectiva, existe $f^{-1} : S \to R$ tal que $f \circ f^{-1} = 1d_S$ y $f^{-1} \circ f = 1d_R$. Por demostrar que $f^{-1} : S \to R$ es un morfismo de anillos.
Para $s_1, s_2 \in S$, existen $r_1, r_2 \in R$ tal que $f(r_1) = s_1$ y $f(r_2) = s_2$ así,

1. $$\begin{aligned}f^{-1}(s_1 + s_2) &= f^{-1}(f(r_1) + f(r_2)) \\ &= f^{-1}(f(r_1 + r_2)) \\ &= r_1 + r_2\end{aligned}$$

$$= f^{-1}(s_1) + f^{-1}(s_2).$$

2. $$\begin{aligned} f^{-1}(s_1 s_2) &= f^{-1}(f(r_1)f(r_2)) \\ &= f^{-1}(f(r_1 r_2)) \\ &= r_1 r_2 \\ &= f^{-1}(s_1)f^{-1}(s_2). \end{aligned}$$

3. $f^{-1}(1_S) = f^{-1}(f(1_R) = 1_R.$

□

Ejemplo 3.10. *Cuando un elemento en un anillo R fue identificado con un polinomio constante. Esto es r fue identificado con $(r, 0, 0, ...)$, implicamos que R es un subanillo de $R[x]$. El subconjunto $R' = \{(r, 0, 0, ...) \mid r \in R\}$ es un subanillo de $R[x]$, además la función $f : R \to R'$ definido por $f(r) = (r, 0, 0, ...)$ es un isomorfismo.*

Demostración. Veamos que $R' = \{(r, 0, 0, ...) \mid r \in R\}$ es un subanillo de $R[x]$

1. $\bar{1} = (1, 0, ...) \in R'$.
2. Si $(r_1, 0, ...), (r_2, 0, ...) \in R'$, entonces $(r_1, 0, ...) - (r_2, 0, ...) = (r_1 - r_2, 0, ...) \in R'$.
3. Si $(r_1, 0, ...), (r_2, 0, ...) \in R'$, entonces $(r_1, 0, ...)(r_2, 0, ...) = (r_1 r_2, 0, ...) \in R'$.

Veamos que $f : R \to R'$ tal que $f(a) = (a, 0, ...)$ es morfismos de anillos.

1. $f(1) = (1, 0, ...)$.
2. $f(a + b) = (a + b, 0, ...) = (a, 0, ...) + (b, 0, ...) = f(a) + f(b)$.
3. $f(ab) = (ab, 0, ...) = (a, 0, ...)(b, 0, ...) = f(a)f(b)$.

Supongamos que $(a, 0, ...) = (a', 0, ...)$ pero esto significa que $a = a'$. Por lo tanto f es inyectiva. Y claramente es sobreyectiva. Por lo tanto $R \cong R' \subset R[x]$. □

Proposición 3.11. *Sean A, R anillos. Si $\varphi : A \to R$ es un morfismo de anillos, entonces $\varphi^* : A[x] \to R[x]$ dado por $\sum a_i x^i \mapsto \sum \varphi(a_i)x^i$ es un morfismo de anillos.*

Demostración. Notemos que φ es un momorfismo entre A y R como grupos aditivos. Por lo tanto $\varphi(0_A) = 0_R$.

1. Probemos que $\varphi^*((1_A, 0, 0, ...)) = (1_R, 0, 0, ...)$. Así $\varphi^*((1_A, 0_A, 0_A, ...)) = (\varphi(1_A), \varphi(0_A), \varphi(0_A) = (1_R, 0_R, 0_R, ...)$.

2. $\varphi^*((a_0, a_1, ..., a_n, 0, 0, ..)+(b_0, b_1, ..., b_n, 0, 0, ...) = \varphi^*((a_0+b_0, a_1+b_1, ..., a_n+b_n, 0, ...)) = (\varphi(a_0+b_0), \varphi(a_1+b_1), ..., \varphi(a_n+b_n), \varphi(0), \varphi(0), ...) = (\varphi(a_0)+\varphi(b_0), ..., \varphi(a_n)+\varphi(b_n), 0_R, 0_R, ...) = (\varphi(a_0), \varphi(a_1), ..., \varphi(a_n), 0_R, 0_R, ...)+(\varphi(b_0), \varphi(b_1), ..., \varphi(b_n), 0_R, 0_r, ...)$.

3. Sea $\sigma = (a_0, a_1, ...,)$ y $\tau = (b_1, b_2, ...,)$. Entonces el coeficiente k-ésimo de $\varphi^*(\sigma\tau)$ es $\varphi(\sum_{i+j=k} a_i b_i) = \sum_{i+j=k} \varphi(a_i)\varphi(b_i)$. Por otro lado

$$\varphi^*(\sigma) = (\varphi(a_0), \varphi(a_1), ...)$$
$$\varphi^*(\tau) = (\varphi(b_0), \varphi(b_1), ...)$$

   Así el coeficiente $k-$ésimo de $\varphi^*(\sigma)\varphi^*(\tau)$ es $\sum_{i+j=k} \varphi(a_i)\varphi(b_i)$.

□

Definición 3.12. *Sea $f : R \to A$ un momorfismo de anillos,*

1. *El núcleo de f está definido por* $\boldsymbol{ker\ f} = \{r \in R \mid f(r) = 0\}$.

2. *La imagen de f está definida por* $\boldsymbol{im\ f} = \{a \in A \mid f(r) = a$ *para algún* $r \in R\}$.

Definición 3.13. *Un ideal en un anillo R es un subconjunto I de R tal que:*

1. $0 \in I$.

2. *Si $a, b \in I$, entonces $a + b \in I$.*

3. *Si $a \in I$ y $r \in R$, entonces $ra \in I$.*

Denotaremos a los ideales por $I \leq R$. El anillo R y el subconjunto que consiste únicamente del elemento 0, el cual denotaremos por $\{0\}$ son siempre ideales del anillo R. Un ideal $I \neq R$ es llamado un ***ideal propio****. A partir de ahora denotaremos a los ideales propios por $I < R$.*

Ejemplo 3.14. *Si $b_1, b_2, ..., b_n \in R$, entonces el conjunto de todas las combinaciones lineales $I = \{r_1b_1 + r_2b_2 + ... + r_nb_n \mid r_i \in R, \forall i \in \{1, 2, ...n\}\}$ es un ideal de R. Escribimos $I = (b_1, ..., b_n)$ en este caso y decimos que I es el* ***ideal generado*** *por $b_1, b_2, ..., b_n$. En particular, si $n = 1$, entonces $I = (b) = \{rb \mid r \in R\}$ es un ideal de R que consiste de todos los múltiplos de b, y este es llamado el* ***ideal principal*** *generado por b. Notemos que R y $\{0\}$ son ideales principales.*

Proposición 3.15. *Un ideal propio de R no contiene unidades.*

Demostración. Supongamos $I < R$ y $u \in I$, unidad, luego existe $v \in R$ tal que $uv = 1$, contradicción, ya que 1 no está en I. □

Proposición 3.16. *Sea $f : R \to A$ un morfismo de anillos, entonces $kerf$ es un ideal de R y imf es un subanillo de A.*

Demostración. Veamos que ker f es un ideal de R.

1. $f(0) = 0$, así $0 \in ker(f)$.
2. Sean $a, b \in kerf$, $f(a + b) = f(a) + f(b) = 0$, así $a + b \in kerf$.
3. $a \in kerf$ y $r \in R$, $f(ra) = f(r)f(a) = 0$, así $ra \in kerf$.

Veamos que imf es un subanillo de A.

1. $f(1) = 1$, así $1 \in imf$.
2. Si $a, b \in imf$, entonces existen $r_1, r_2 \in R$ tales que $f(r_1) = a$ y $f(r_2) = b$, por lo que, $a - b = f(r_1) - f(r_2) = f(r_1 - r_2) \in imf$.
3. Sean $a, b \in imf$, $ab = f(r_1)f(r_2) = f(r_1r_2) \in imf$.

□

Proposición 3.17. *Sea $f : R \to A$ un morfismo de anillos. Entonces f es inyectivo si y sólo si $kerf = \{0\}$.*

Demostración. $\Rightarrow$] Sea $a \in kerf$ entonces $f(a) = 0$ y $f(0) = 0$ ya que $0 + f(0) = f(0 + 0) = f(0) + f(0)$, así $f(0) = f(a)$ y como f es inyectiva, $a = 0$. Por lo tanto $kerf = \{0\}$.
$\Leftarrow$] $f(a) = f(b)$ si y sólo si $f(a) - f(b) = 0$, luego $f(a - b) = 0$ así $a - b \in kerf = \{0\}$. Por lo tanto $a - b = 0$ implica $a = b$. □

Proposición 3.18. *Sea $f : R \to S$ un morfismo de anillos suprayectivo. Si I es un ideal de R, entonces $f(I)$ es un ideal de S.*

Demostración. Basta demostrar que para todo $s \in S$ y $x \in f(I)$ se cumple que $sx \in f(I)$. Sea $x \in f(I)$ entones $x = f(a)$ para algún $a \in I$ y sea $s \in S$, entonces existe $r \in R$ tal que $f(r) = s$, así $sx = f(r)f(a) = f(ra)$, con $ra \in I$. Por lo tanto $sx = f(ra) \in f(I)$. □

Proposición 3.19. *Sea $f : R \to S$ morfismo de anillos. Si J es un ideal de S, entonces $f^{-1}(J)$ es un ideal de R.*

Demostración. Basta demostrar que para todo $r \in R$ y para todo $x \in f^{-1}(J)$, $rx \in f^{-1}(J)$.
Sean $r \in R$ y $x \in f^{-1}(J)$, así $f(x) \in J$. Entonces $f(rx) = f(r) \cdot f(x) \in J$. Por lo tanto $rx \in f^{-1}(J)$. □

Proposición 3.20. *Si $f : R \to S$ es un morfismo de anillos, tal que $I \subseteq J \subseteq R$, entonces $f(I) \subseteq f(J)$*

Demostración. Si $x \in f(I)$, entonces $x = f(a)$ para algún $a \in I$, como $I \subseteq J$ entonces $x = f(a) \in f(J)$. □

Proposición 3.21. *Si $f : R \to S$ es un morfismo de anillos, tal que si $I' \subseteq K' \subseteq S$, entonces $f^{-1}(I') \subseteq f^{-1}(K')$.*

Demostración. Si $x \in f^{-1}(I')$, entonces $f(x) \in I' \subseteq K'$. Por lo tanto $x \in f^{-1}(K')$. □

Proposición 3.22. *Sea $f : R \to S$ es un morfismo de anillos y I un ideal de R. Entonces $f^{-1}f(I) = Kerf + I$.*

Demostración. $\subseteq$] Sea $y \in f^{-1}f(I)$, entonces $f(y) \in f(I)$ así $f(y) = f(a)$, para algún $a \in I$, entonces $f(y-a) = 0_S$, de modo que $(y-a) \in Ker(f)$ así $y = a + (y-a)$ con $a \in I$ y $(y-a) \in Ker(f)$. Por lo tanto $y \in I + Ker(f)$.
$\supseteq$] Sabemos que $0_S \in f(I)$, entonces $f^{-1}(0_S) \subseteq f^{-1}f(I)$. Por lo tanto $kerf \subseteq f^{-1}(f(I))$.
Si $x \in I$, $f(x) \in f(I)$, entonces $x \in f^{-1}(f(I))$, así $I \subseteq f^{-1}(f(I))$. Entonces $ker(f) + I \subseteq f^{-1}f(I)$. □

Más adelante cuando contemos con la estructura de anillo cociente, podremos seguir mostrando propiedades de los morfismos sobre dicha estructura.

3.1. Dominio de Ideales Principales

Es común en álgebra abstraer. Debido a que el anillo de los enteros es un dominio de ideales principales y que de este hecho surja una teoría de divisibilidad no es ninguna casualidad. En está corta sección se trabaja con este concepto solamente lo necesario.

Definición 3.23. *Sea R un Dominio Entero. R es un* ***Dominio de Ideales Principales*** *(D.I.P) si para cada ideal I de R, existe un elemento $a \in R$ tal que $\langle a \rangle = I$.*

Proposición 3.24. *Sea $\mathbb{F}$ un campo entonces $\mathbb{F}[x]$ es un D.I.P. Además para cada $I \subseteq \mathbb{F}[x]$, existe $f(x) \in \mathbb{F}[x]$ mónico tal que $\langle f(x) \rangle = I$*

Demostración. Si $\mathbb{F}$ un campo entonces $\mathbb{F}[x]$ es un ejemplo de un dominio euclidiano. En la Proposición 3.29 probaremos que todo ideal en un anillo euclidiano es un ideal princial. □

Definición 3.25. *Sea R un anillo. Sean $a, b, \delta, \gamma \in R$. δ es un máximo común divisor si:*

1. *$\delta \mid a$ y $\delta \mid b$.*
2. *Si γ es otro común divisor de a y b, entonces $\gamma \mid \delta$.*

Proposición 3.26. *Sea R un D.I.P*

1. *Para todo $a, b \in R$ existe un m.c.d, δ, el cual es combinación lineal de a y b :*

$$\delta = \sigma a + \tau b,$$

   para algunos $\sigma, \tau \in R$.

2. *Si p es irreducible y $p \mid ab$, entonces $p \mid a$ ó $p \mid b$.*

Demostración. 1. Si $a = b = 0$, entonces $(0, 0) = 0 = 0a + 0b$. Consideremos el conjunto J de todas las combinaciones lineales

$$J = \{\sigma a + \tau b \mid \sigma, \tau \in R\}.$$

Ahora a y b están en J, ya que podemos tomar $\sigma = 1$ y $\tau = 0$ o viceversa. Es necesario ver que J es un ideal:

a) $0 \in I$, ya que si tomamos $\sigma = \tau = 0$, entonces $\sigma a + \tau b = 0$.

b) Si $x, y \in J$, entonces $x = \sigma' a + \tau' b$, $y = \sigma'' a + \tau'' b$, luego

$$x + y = (\sigma' a + \tau' b) + (\sigma'' a + \tau'' b) = (\sigma' + \sigma'')a + (\tau' + \tau'')b \in J.$$

c) Si $x \in J$, $r \in R$, entonces $x = \sigma a + \tau b$, luego $rx = r(\sigma a + \tau b) = (r\sigma)a + (r\tau)b \in J$.

Y así existe $\delta \in J$ tal que $J = \langle \delta \rangle$, ya que R es un D.I.P. Afirmamos que δ es un m.c.d de a y b. Como $a \in J = \langle \delta \rangle$, tenemos que $a = p\delta$ para algún $p \in R$, esto es, δ es un divisor de a, similarmente, δ es un divisor de b, y así δ es un común divisor de a y b.
Ya que $\delta \in I$, es una combinación lineal de a y b, entonces existen $\sigma, \tau \in R$ con

$$\delta = \sigma a + \tau b.$$

Finalmente, si γ es otro común divisor de a y b, entonces $a = \gamma a'$ y $b = \gamma b'$, luego $\delta = \sigma a + \tau b = \gamma(\sigma a' + \tau b')$, así γ divide δ. Y concluimos que δ es un m.c.d.

2. Si $p \mid a$ se ha terminado. Si $p \nmid a$ entonces $(p, a) = 1$, entonces existen $s, t \in R$ tal que $sp + ta = 1$, luego multiplicando por b en ambos lados, $spb + tab = b$. Como $p \mid ab$ entonces existe $h \in R$ tal que $ab = hp$, así $b = (sb + th)p$. Esto implica que $p \mid b$.

□

Definición 3.27. *Un **Dominio Euclidiano** es un Dominio Entero R, que está equipado con una función*

$$\partial : R \backslash \{0\} \longrightarrow \mathbb{N}$$

llamado una función grado, tal que.

1. $\partial(f) \leq \partial(f \cdot g)$ *para todo* $f, g \in R$ *con* $f, g \neq 0$.

2. *Para todo* $f, g \in R$ *con* $f \neq 0$, *existe* $q, r \in R$ *con*

$$g = qf + r,$$

donde $r = 0$ *ó* $\partial(r) < \partial(f)$.

Ejemplo 3.28. *Si* K *es un campo, el dominio* $K[x]$ *es un anillo euclidiano con función grado el grado usual de un polinomio distinto de cero. En* $K[x]$*, tenemos,*

$$\begin{aligned}\partial(f \cdot g) = gr(f \cdot g) &= gr(f) + gr(g) \\ &= \partial(f) + \partial(g)\end{aligned}$$

Proposición 3.29. *Todo Dominio Euclidiano es un D.I.P.*

Demostración. Sea R un dominio euclidiano e I un ideal de R. Si $I = \{0\}$, entonces $I = \langle 0 \rangle$ es principal. Si $I \neq \{0\}$, por el Principio del Buen Orden, en el conjunto de todos los grados de elementos distintos de cero en I existe un elemento de menor grado $0 \neq a \in I$ con $\partial(a) = n$. Claramente $\langle a \rangle \subseteq I$. Sea $b \in I$, puesto que $a \neq 0$, existen $q, r \in R$ tal que $b = qa + r$ con $r = 0$ ó $\partial(r) < \partial(a)$. Supongamos que $r \neq 0$, entonces $r = b - qa \in I$, lo que es una contradicción, ya que r tiene grado menor que a. Por lo tanto $r = 0$, así $b = qa$ y así $b \in \langle a \rangle$. Así $I \subseteq \langle a \rangle$. Por lo tanto $I = \langle a \rangle$. □

Capítulo 4

Espacios Vectoriales

Un código lineal de longitud n sobre el campo finito $\mathbb{F}_q$ no es más que un subespacio del espacio vectorial $\mathbb{F}_q^n$. Ya que los códigos lineales son espacios vectoriales, será importante recordar nuevamente las propiedades más importantes que poseen estós,

Definición 4.1. *Sea $\mathbb{F}_q$ un campo finito. Un espacio vectorial es un grupo aditivo abeliano $(V, +)$ con un producto escalar*

$$\begin{array}{c} \mathbb{F}_q \times V \longrightarrow V \\ (k, v) \longrightarrow kv \end{array}$$

tal que:

I. $k(v_1 + v_2) = kv_1 + kv_2$ *para todo* $k \in \mathbb{F}_q$, $v_1, v_2 \in V$.

II. $(k_1 + k_2)v = k_1v + k_2v$ *para todo* $k_1, k_2 \in \mathbb{F}_q$ *y* $v \in V$.

III. $(k_1k_2)v = k_1(k_2v)$ *para todo* $k_1, k_2 \in \mathbb{F}_q$ *y* $v \in V$.

IV. $1_kv = v$ *para todo* $v \in V$.

A los elementos de V los llamamos vectores y a los elementos de $\mathbb{F}_q$ los llamamos escalares.

Ejemplo 4.2. *Sea $\mathbb{F}$ un campo y X un conjunto. Entonces*

$$\mathbb{F}^X = \{f : X \to \mathbb{F} \mid f \text{ es una función } \}.$$

Sea define la suma de funciones de la manera usual, y el producto de un elemento de $\mathbb{F}$ por una función también de la manera usual.

1. $f\hat{+}g : X \to \mathbb{F}$ *es la función tal que* $(f\hat{+}g)(x) = f(x) + g(x)$.
2. $c \cdot f : X \to \mathbb{F}$ *es la función tal que* $(c \cdot f)(x) = c(f(x))$.

Entonces $(\mathbb{F}^X, \hat{+}, \hat{0}, \cdot : \mathbb{F} \times \mathbb{F}^X \to \mathbb{F}^X)$ *es un espacio vectorial.*

Demostración. Veamos que $(\mathbb{F}^X, \hat{+}, \hat{0})$ es un grupo abeliano.

1. $$\begin{aligned}((f\hat{+}g)\hat{+}h)(x) &= (f+g)(x) + h(x) \\ &= (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) \\ &= f(x) + (g+h)(x) \\ &= (f\hat{+}(g\hat{+}h))(x).\end{aligned}$$

2. Definimos $\hat{0} : X \to \mathbb{F}$ como $\hat{0}(x) = 0$ y es tal que

$$(\hat{0}\hat{+}f)(x) = \hat{0}(x) + f(x) = 0 + f(x) = f(x) + 0 = (f\hat{+}\hat{0})(x).$$

3. Para toda $f \in \mathbb{F}^X$, existe $-f \in \mathbb{F}^X$, definida por $-f : X \to \mathbb{F}$ tal que $x \mapsto -f(x)$, por lo tanto $f\hat{+}(-f))(x) = f(x)+(-f(x)) = f(x)-f(x) = 0$

4. $(f\hat{+}g)(x) = f(x) + g(x) = g(x) + f(x) = (g\hat{+}f)(x)$.

Veamos que se cumplen las propiedades del producto por escalares:

1. $(1 \cdot f)(x) = 1 \cdot f(x) = f(x)$, para toda $x \in X$.
2. $[(cd) \cdot f](x) = (cd)f(x) = c(df(x)) = c((df)(x)) = (c \cdot (d \cdot f))(x)$, para toda $x \in X$.
3. $[(c+d)(f)](x) = (c+d)(f(x)) = cf(x) + df(x) = (cf)(x) + (df)(x) = (cf + df)(x)$, para toda $x \in X$.
4. $[c \cdot (f\hat{+}g)](x) = c[(f\hat{+}g)(x)] = c[f(x) + g(x)] = c(f(x)) + c(g(x)) = (c \cdot f)(x) + (c \cdot g)(x)$, para todo $x \in X$.

Por lo tanto $(\mathbb{F}^X, \hat{+}, \hat{0}, \cdot : \mathbb{F} \times \mathbb{F}^X \to \mathbb{F}^X)$ es un espacio vectorial. □

Ejemplo 4.3. *Sea* $\mathbb{F}_q$ *un campo, y* $\mathbb{F}_q^n$ *el conjunto*

$$\mathbb{F}_q^n = \{(x_1, ..., x_n) : x_i \in \mathbb{F}_q\}$$

donde están definidas las siguientes operaciones

$$(a_1, a_2, ..., a_n)+(b_1, b_2, ..., b_n)= (a_1 + b_1, a_2 + b_2, ..., a_n + b_n)$$
$$\alpha(a_1, a_2, ..., a_n)= (\alpha a_1, \alpha a_2, ..., \alpha a_n)$$

Estas operaciones hacen que $\mathbb{F}_q^n$ sea un espacio vectorial sobre $\mathbb{F}_q$.

Definición 4.4. *Sean $\mathbb{F}$ un campo y $X = \{1, 2, ..., n\} \times \{1, 2, ..., m\}$, un elemento A en $\mathbb{F}^X$ se llama una matriz de $n \times m$ con coeficientes en $\mathbb{F}$ y la denotamos como $A \in \mathbb{M}_{n\times m}(\mathbb{F})$ y notemos que es un elemento de $\mathbb{F}^X$.*
Por costumbre, uno escribe $A_{i,j}$ en lugar de escribir $A(i, j)$. También por costumbre, uno suele escribir una matriz A en la forma de un arreglo rectangular:

$$\begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & ... & A_{1,m} \\ A_{2,1} & A_{2,2} & A_{2,3} & ... & A_{2,m} \\ \vdots & & & & \\ A_{n,1} & A_{n,2} & A_{n,3} & ... & A_{n,m} \end{pmatrix}$$

Definición 4.5. *Si $A \in \mathbb{M}_{n\times m}(\mathbb{F})$, su transpuesta es la matriz $A^T \in \mathbb{M}_{m\times n}(\mathbb{F})$ tal que*

$$A^T_{i,j} = A_{j,i}.$$

Definición 4.6. *Sea $\mathbb{F}_q$ campo y V un espacio vectorial. Un subconjunto $U \subseteq V$ es un subespacio vectorial de V si :*

1. *$0 \in U$.*
2. *Si $u_1, u_2 \in U$ entonces $u_1 + u_2 \in U$.*
3. *Si $u \in U$ y $r \in \mathbb{F}_q$ entonces $ru \in U$.*

Ejemplo 4.7. *Usando la notación del Ejemplo 4.3 con $q = 3$ y $n = 3$, $C = \{(0, 0, 0), (0, 1, 2), (0, 2, 1)\}$ es subespacio de $\mathbb{F}_3^3$.*

Proposición 4.8. *Un subconjunto no vacío C de un espacio vectorial V sobre $\mathbb{F}$ es un subespacio si y sólo si la siguiente condición se satistace:*

$$\text{Si } x, y \in C \text{ y } \lambda \in \mathbb{F}_q, \text{ entonces } \lambda x + y \in C.$$

Demostración. Supongamos que C es un conjunto no vacío de V tal que $\lambda x+y$ pertenece a C para todos los vectores $x, y \in C$ y todos los escalares $\lambda \in \mathbb{F}$. Ya que C es no vacío, existe un vector p en C y por lo tanto $(-1)p+p = 0$ esta en C. Entonces si x es cualquier vector en C y λ cualquier escalar, el vector $\lambda x=\lambda x +0$ está en C. En particular, $(-1)x = -x$ está en C. Finalmente si x e y están en C, entonces $x + y= 1x + y$ está en C. Por lo tanto C es un subespacio vectorial de V. Recíprocamente, si C es un subespacio de V, x e y en C y λ un escalar, entonces $\lambda x + y$ está en C. □

Definición 4.9. *Sea V un espacio vectorial sobre $\mathbb{F}$. Sea $X = \{v_1, ..., v_n\}$ un subconjunto no vacío de V, una combinación lineal de X es un vector de la forma*

$\sum_{i=1}^{n} r_i v_i = r_1 v_1 + ... + r_n v_n \in V$ *donde* $r_i \in \mathbb{F}_q$ *para toda* $i \in \{1, 2, ..., n\}$.

Denotamos al conjunto de todas las combinaciones lineales de $v_1, ..., v_n$ como

$$\langle v_1, ..., v_n \rangle = \{\textstyle\sum_{i=1}^{n} r_i v_i \mid r_i \in \mathbb{F}_q\} = \langle X \rangle.$$

Note que $\langle X \rangle$ es el menor subespacio que contiene a X. Por lo tanto $\langle X \rangle$ es el subespacio generado por X. Además si $X = \emptyset$ es claro que $\{0\}$ es el menor subespacio que contiene al vacío. Por lo que $\langle \emptyset \rangle = \{0\}$.

Ejemplo 4.10. *Sea $A \in \mathbb{M}_{n\times m}(\mathbb{F})$ el **espacio columna** de A es el conjunto de aquellos vectores de $\mathbb{F}^n$ que se pueden expresar como combinaciones lineales de las m columnas de la matriz A. Así el espacio columna consiste de aquellos vectores de la forma*

$$x_1 a_1 + x_2 a_2 + ... + x_m a_m.$$

Donde los $x_i \in \mathbb{F}$ y los a_i son las columnas de A.
*Sea $A \in \mathbb{M}_{n\times m}(\mathbb{F})$, el **espacio fila** de A es el conjunto de aquellos vectores de $\mathbb{F}^m$ que se pueden expresar como combinaciones lineales de los n renglones de la matriz A. Así el espacio fila consiste de aquellos vectores de la forma*

$$x_1 a_1 + x_2 a_2 + ... + x_n a_n.$$

Donde los $x_i \in \mathbb{F}$ y los a_i son las filas de A.

Definición 4.11. *Un subconjunto no vacío $X = \{v_1, ..., v_n\}$ de vectores en V, es **linealmente dependiente** si existen $a_1, ..., a_n \in K$ no todos ceros tal que*

$$\sum_{i=1}^{n} a_i v_i = 0.$$

*X es **linealmente independiente** si la relación*

$$\sum_{i=1}^{n} a_i v_i = 0.$$

sólo se satisface si $a_i = 0$ para cada $i \in \{1, ...n\}$.

Ejemplo 4.12. *1. Cualquier conjunto S que contenga al 0 es linealmente dependiente.*

2. Para cualquier $\mathbb{F}_q$, el conjunto $\{(0,0,0,1),(0,0,1,0),(0,1,0,1)\}$ es linealmente independiente.

Definición 4.13. *Una base de un espacio vectorial V es un conjunto X linealmente independiente tal que $V = \langle X \rangle$.*

Proposición 4.14. *Todo espacio vectorial distinto del cero tiene una base.*

Demostración. Sea V un espacio vectorial distinto del cero, y sea $S = \{L \subseteq V \mid L \text{ es linealmente independiente}\}$. Un vector $0 \neq v \in V$ es linealmente independiente, así $\{v\} \in S$, por lo tanto $S \neq \emptyset$. Para dos conjuntos linealmente independientes, afirmamos $L \leq L'$ si $L \subset L'$. Este es el orden parcial sobre S dado por la inclusión. Además cualquier subconjunto de un conjunto linealmente independiente es también linealmente independiente, así si $L \in S$ entonces cualquier subconjunto de L está también en S. Asumamos que $\mathcal{C} = \{L_i\}_{i \in I}$ es una cadena de S. Esto es, todo L_i es un conjunto linealmente independiente en V y para todo L_i y L_j en $\mathcal{C}$ tenemos que $L_i \subset L_j$ ó $L_j \subset L_i$. Afirmamos que

$$L = \bigcup_{i \in I} L_i,$$

es un cota superior en S. Necesitamos mostrar que L es un conjunto linealmente independiente, es decir $L \in S$. Escogemos un conjunto finito de vectores $v_1, ...v_n \in L$. Así cada v_k está en algún L_i, digamos $v_1 \in L_{i_1}, ..., v_n \in L_{i_n}$. Ya que los L_i están totalmente ordenados, uno de los conjuntos $L_{i_1}, ..., L_{i_n}$ contiene a los demás. Esto significa que $v_1, ..., v_n$ están todos en algún L_i, y así estos son linealmente independientes. Por el Lema de Zorn, S contiene un

elemento máximo, es decir, existe un conjunto B linealmente independiente en V que no está contenido en ningún conjunto linealmente independiente más grande en V. Vamos a demostrar que B genera a V, es decir es una base. Sea $\mathcal{W}$ el subespacio generado por B. Esto significa que $\mathcal{W}$ es el conjunto de todas las combinaciones lineales finitas $\sum_{i=1}^{k} c_i v_i$ con $k \geq 1$, $c_i \in \mathbb{F}$, y $v_i \in B$. Si B no genera V entonces $\mathcal{W} \neq V$, así podemos escoger $v \in V$ con $v \notin \mathcal{W}$. Entonces B es un subconjunto propio de $B \cup \{v\}$. Vamos a demostrar que $B \cup \{v\}$ es linealmente independiente, lo que contradice el hecho que B es máximo y por lo tanto $\mathcal{W} = V$.
Para probar que $B \cup \{v\}$ es linealmente independiente, asumamos lo contrario, es decir, que existe una expresión

$$\sum_{i=1}^{k} c_i v_i = 0$$

donde los coeficientes no son todos cero y los v_i son tomados de $B \cup \{v\}$. Ya que los elementos de B son linealmente independientes, uno de los v_i con un coeficiente distinto del cero tiene que ser v. Sin perdida de generalidad supongamos que $v_k = v$, así $c_k \neq 0$. Debemos tener que $k \geq 2$. ya que de otra manera $c_1 v = 0$, lo cual es una contradicción ya que $v \neq 0$ y el coeficiente de v es distinto de cero. Entonces

$$c_k v = -\sum_{i=1}^{k-1} c_i v_i.$$

Multiplicando en ambos lados por $1/c_k$,

$$v = \sum_{i=1}^{k-1} \left(-\frac{c_i}{c_k} \right) v_i,$$

lo que muestra que $v \in \mathcal{W}$. Pero $v \notin \mathcal{W}$. Por lo tanto $B \cup \{v\}$ es un conjunto linealmente independiente. □

Proposición 4.15. *Sean $u_1, ..., u_n$ elementos en un espacio vectorial V, y sea $v_1, ..., v_m \in \langle u_1, ..., u_n \rangle$. Si $m > n$ entonces $\{v_1, ..., v_m\}$ es un conjunto linealmente dependiente.*

Demostración. La demostración se hará por inducción sobre $n \geq 1$.
Si $n = 1$, entonces $m > 1$, entoces existen al menos dos vectores $v_1, v_2 \in \langle u_1 \rangle$. Si $u_1 = 0$, entonces $v_1 = 0$ y así $\{v_1, ..., v_m\}$ es un conjunto linealmente dependiente. Supongamos que $u_1 \neq 0$. Podemos considerar $v_1 \neq 0 \neq v_2$.

Ahora como $v_1, v_2 \in \langle u_1 \rangle$, entonces existen $a, b \in K$ tal que $v_1 = au_1$ y $v_2 = bu_1$, entonces $a \neq 0$, implica $u_1 = a^{-1}v_1$, así $v_2 = ba^{-1}v_1$, por lo tanto $1v_2 - ba^{-1}v_1 = 0$ es una combinación lineal de elementos de $\{v_1, ..., v_m\}$ no trivial. Por lo tanto el conjunto $\{v_1, ..., v_m\}$ es linealmente dependiente.
Para $n > 1$, existen ecuaciones, para $i = 1, ...m$,

$$v_i = a_{i1}u_1 + ... + a_{in}u_n.$$

Podemos asumir que algún $a_{i1} \neq 0$, ya que si $a_{j1} = 0$ para $j \in \{1, ..., m\}$, entonces $v_1, ..., v_m \in \langle u_2, ..., u_n \rangle$, y por hipótesis inductiva $\{v_1, ..., v_m\}$ es linealmente dependiente. Supongamos sin pérdida de generalidad que $a_{11} \neq 0$. Para cada $i \geq 2$, definimos

$$v_i' = v_i - a_{i1}a_{11}^{-1}v_1 \in \langle u_2, ..., u_n \rangle$$

Ya que $m - 1 > n - 1$, por hipótesis inductiva, existe escalares $b_2, ..., b_m$ no todos cero, con

$$b_2v_2' + ... + b_mv_m' = 0.$$

Reescribiendo la ecuación usando la igualdad de v_i':

$$\left(-\sum_{i \geq 2} b_i a_{i1} a_{11}^{-1} \right) v_1 + b_2v_2 + ... + b_mv_m = 0$$

Notemos que no todos los coeficientes son 0, así el conjunto $\{v_1, ..., v_m\}$ es linealmente dependiente. □

Proposición 4.16. *Sean $X = \{x_1, ..., x_n\}$ y $Y = \{y_1, ..., y_m\}$ dos bases de un espacio vectorial V sobre $\mathbb{F}_q$ entonces $m = n$.*

Demostración. Supongamos que $m > n$ entonces $y_1, ..., y_m \in \langle x_1, ..., x_n \rangle = V$. Por ser X una base, entonces $\{y_1, ..., y_m\}$ es linealmente dependiente, contradicción.
Si $m < n$, entonces $x_1, .., x_n \in \langle y_1, ..., y_m \rangle = V$. Por ser Y una base, entonces $\{x_1, .., x_n\}$ es linealmente dependiente, contradicción. Por lo tanto $m = n$. □

Definición 4.17. *Un espacio vectorial V es llamado de **dimensión finita** si tiene una base que consiste de un número finito de vectores. El único número de vectores en cada base para V es llamada la dimensión de V y denotada por $dim(V)$.*

Definición 4.18. *La dimensión del espacio fila de la matriz A es llamado el **rango** de la matriz A*

Ejemplo 4.19. *1. El espacio vectorial* $\{0\}$ *tiene dimensión cero.*

2. El espacio vectorial $\mathbb{F}_q^n$ *tiene dimensión* n*.*

Proposición 4.20. *Sea* $X = \{v_1, ...v_n\}$ *un conjunto de vectores de* V*.* X *es base de* V *si y sólo si todo vector en* V *tiene una única expresión como combinación lineal de los elementos de* X*.*

Demostración. $\Rightarrow$] Sea X es base de V y $u \in V$, entonces existen $r_i \in \mathbb{F}_q$ tal que

$$u = \sum_{i=1}^{n} r_i v_i$$

y supongamos que u también se puede escribir como

$$u = \sum_{i=1}^{n} s_i v_i \text{ y con } s_i \in \mathbb{F}_q.$$

entonces

$$\sum_{i=1}^{n} r_i v_i - \sum_{i=1}^{n} s_i v_i = \sum_{i=1}^{n} (r_i - s_i) v_i = 0$$

puesto que X es linealmente independiente, entonces $r_i - s_i = 0$, de modo que $r_i = s_i$ para todo $1 \leq i \leq n$.
$\Leftarrow$] Todo vector es combinación lineal de los elementos de X, así $\langle X \rangle = V$. Por otro lado si

$$\sum_{i=1}^{n} r_i v_i = 0 = \sum_{i=1}^{n} 0 v_i$$

y dado que la expresión es única, tenemos que $a_i = 0$ para todo $1 \leq i \leq n$, por lo tanto X es linealmente independiente y así X es base de V.

□

Proposición 4.21. *Cualesquiera vectores linealmente independientes* $a_1, ..., a_m$ *con* $m \leq k$*, en un espacio vectorial de dimensión* k*, forma parte de una base* $a_1, ..., a_m, b_{m+1}, ..., b_k$ *de ese espacio vectorial.*

Demostración. Se hará por inducción sobre m.

1. Para $m = 1$. Dado un vector a el cual es linealmente independiente, es decir, distinto de cero, en un espacio lineal V, escojemos una base arbitraria de k vectores en V, $b_1, ..., b_k$. Podemos expresar a como una combinación lineal $a = \sum_{i=1}^{k} t_i b_i$ Alguno de los coeficientes t_i es distinto de cero, ya que $a \neq 0$. Supongamos por ejemplo, que $t_1 \neq 0$. Entonces mostraremos que $a, b_2, ..., b_k$ forma una base de V. En efecto:

   a. Los vectores $a, b_2, ..., b_k$ generan el espacio V porque cada uno de los vectores b_i que generan V, es una combinación de esos vectores. Esto es obvio para $i = 2, ..., k$ y para $i = 1$, tenemos

   $$b_1 = t_1^{-1} a - \sum_{i=2}^{k} (t_1^{-1} t_i) b_i.$$

   b. Para probar la independencia lineal, consideremos una combinación lineal

   $$s_1 a + s_2 b_2 + ... + s_k b_k = 0.$$

   Ya que $a = \sum_{i=1}^{k} t_i b_i$, tenemos

   $$s_1 t_1 b_1 + (s_2 + s_1 t_2) b_2 + ... + (s_k + s_1 t_k) b_k = 0.$$

   De la independencia lineal de $b_1, ..., b_k$, concluimos que $s_1 t_1 = 0$ (así, $s_1 = 0$ porque $t_1 \neq 0$ por hipótesis) y como $s_i + s_1 t_i = 0$, implica $s_i = 0$ para $i = 2, ..., k$. De ahí la combinación lineal es trivial.

2. Paso inductivo: Supongamos que se cumple para $a_1, ..., a_{m-1}$ es decir, existe una base $a_1, ..., a_{m-1}, b_m, ..., b_k$. Tenemos que mostrar que la proposición se cumple para $a_1, ..., a_m$. Podemos expresar a_m como una combinación lineal

   $$a_m = t_1 a_1 + ... + t_{m-1} a_{m-1} + s_m b_m + ... + s_k b_k.$$

   Alguno de los coeficientes s_i es distinto de cero (ya que a_m no es una combinación lineal de $a_1, ..., a_{m-1}$). Supongamos,por ejemplo que $s_m \neq 0$. Entonces $a_1, ..., a_{m-1}, a_m, b_{m+1}, ..., b_k$ es una base de V. Esto se sigue, de manera analoga al paso anterior $m = 1$, de el hecho que b_m es una combinación lineal de estos vectores:

   $$b_m = \sum_{i=1}^{m-1} (-s_m^{-1} t_i) a_i + s_m^{-1} a_m + \sum_{i=m+1}^{k} (-s_m^{-1} s_i) b_i.$$

□

Proposición 4.22. *En un espacio vectorial V de dimensión k se tiene las siguientes propiedades:*

1. *Cada k vectores linealmente independientes forma una base.*

2. *k es el número más grande de vectores linealmente independientes en V.*

3. *Todo subespacio de V, excepto V, tiene dimensión más pequeña que k.*

Demostración. 1. Es claro por la Proposición 4.21.

2. Es claro por la Proposición 4.21.

3. Sea K un subespacio lineal de V y sea m el número más grande de vectores linealmente independentes en K. Por 2, sabemos que $m \leq k$. Cualquier colección linealmente independiente $a_1, ..., a_m$ de vectores en K es una base de K. (De hecho, para cada vector $a \neq 0$, la colección $a, a_1, ..., a_m$ es linealmente dependiente, así existe una combinación lineal no trivial $ta + \sum t_i a_i = 0$. La independencia lineal de $a_1, ..., a_m$ implica que $t \neq 0$, y entonces $a = \sum(-t^{-1}t_i)a_i$. De ahí, $a_1, ..., a_m$ genera K.) Si $m = k$, entonces, por 1. $a_1, ..., a_m$ forma una base de V, y así. $V = K$. En otras palabras, si $K \neq V$, entonces $m < k$.

□

Proposición 4.23. *Sea V un espacio vectorial sobre $\mathbb{F}_q$. Si $dim(V) = k$, entonces V tiene q^k elementos.*

Demostración. Si $\{v_1, ..., v_k\}$ es una base para V, entonces

$$V = \{a_1 v_1 + ... + a_k v_k \mid a_1, ..., a_k \in \mathbb{F}_q\}$$

Ya que $|\mathbb{F}_q| = q$, existen exactamente q elecciones para cada de $a_1, ..., a_k$; por lo tanto, V tiene exactamente q^k elementos. □

Veamos una última propiedad de las matrices.

Proposición 4.24. *Toda matriz de rango k tiene k columnas linealmente independientes.*

Demostración. Si una matriz A está en forma escalonada, entonces esta tiene k filas distintas de cero y seleccionamos las k columnas en las que se encuentran los coeficientes principales de las filas. Estas columnas son linealmente independientes porque si los escribimos como filas (de arriba hacía abajo), claramente tenemos una matriz de k filas distintas de cero en forma escalonada. Ya que toda matriz puede ser puesta en forma escalonada por una sucesión de operaciones elementales en las filas. Es suficiente mostrar que ninguna de las operaciones en las filas cambia la independencia lineal de las columnas. Presentamos la demostración para el caso del intercambio de dos filas (los otros dos casos de operaciones elementales en filas son análogos). Consideremos k columnas de la matriz A con número de columnas $j_1, j_2, ..., j_k$. Si las escribimos como filas, obtenemos k vectores $b_{j_1}, ..., b_{j_k}$. Ahora intercambiamos la i-ésima y la i'-ésima filas de la matriz A. Las correspondientes columnas $\bar{b_{j_1}}, ..., \bar{b_{j_k}}$ de la nueva matriz son obtenidos de los vectores originales $b_{j_1}, ..., b_{j_k}$ intercambiando las i-ésima y i'-ésima posiciones. Por lo tanto demostraremos que los vectores $b_{j_1}, ..., b_{j_k}$ son linealmente independientes si y sólo si así lo son los nuevos vectores $\bar{b_{j_1}}, ..., \bar{b_{j_k}}$. Pero ya que, dado escalares $t_1, ..., t_k$ claramente tenemos

$$t_1 b_{j_1} + ... + t_k b_{j_k} = 0 \text{ si y sólo si } t_1 \bar{b_{j_1}} + ... + t_k \bar{b_{j_k}} = 0.$$

□

Definición 4.25. *Sean* $\boldsymbol{v} = (v_1, v_2, ..., v_n)$ *y* $\boldsymbol{w} = (w_1, w_2, ..., w_n) \in \mathbb{F}^n$

1. *El* ***producto escalar*** *o el* ***producto punto*** *de* $\boldsymbol{v}$ *y* $\boldsymbol{w}$ *está definido como*

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_1 w_1 + ... + v_n w_n \in \mathbb{F}.$$

2. *Los dos vectores* $\boldsymbol{v}$ *y* $\boldsymbol{w}$ *se dicen ortogonales si* $\boldsymbol{v} \cdot \boldsymbol{w} = 0$.

3. *Sea* S *un subconjunto no vacío de* $\mathbb{F}_q^n$. *El* ***complemento ortogonal*** $S^\perp$ *de* S *está definido por*

$$S^\perp = \{\boldsymbol{v} \in \mathbb{F}_q^n \mid \boldsymbol{v} \cdot \boldsymbol{s} = 0 \text{ para todo } \boldsymbol{s} \in S\}.$$

Si $S = \emptyset$, *entonces* $S^\perp = \mathbb{F}_q^n$.

Lema 4.26. $S^\perp$ *es un subespacio vectorial de* $\mathbb{F}_q^n$ *para cualquier subconjunto de* $\mathbb{F}_q^n$, *y que* $\langle S \rangle^\perp = S^\perp$.

Demostración. Sean $x, y \in S^{\perp}$ y $\alpha, \beta \in \mathbb{F}_q$. Así $x \cdot s = 0$ y $y \cdot s = 0$ para toda $s \in S$. Además $\alpha(x \cdot s) = 0$ y $\beta(y \cdot s) = 0$, por lo tanto $0 = \alpha(x \cdot s) + \beta(y \cdot s) = (\alpha x) \cdot s + (\beta y) \cdot s = (\alpha x + \beta y) \cdot s$. Por la Proposición 4.8, $S^{\perp}$ es un subespacio vectorial de $\mathbb{F}_q^n$. □

Proposición 4.27. *El complemento ortogonal de un subespacio L de dimensión k del espacio vectorial $\mathbb{F}^n$, tiene dimensión $dim(L^{\perp}) = n - dim(L)$.*

Demostración. Sea $a_1, ..., a_k$ una base de un subespacio lineal L de $\mathbb{F}^n$. Aquellos vectores, escritos como filas, forman una matriz $A \in \mathbb{M}_{k \times n}(\mathbb{F})$. Un vector b esta en el complemento ortogonal de L si y sólo si $Ab^T = 0^T$. Ya que A tiene rango k, tiene k columnas linealmente independientes por la Proposición 4.24. La longitud de las columnas es k, y así por la Proposición 4.22 (3) las columnas generan el espacio lineal $\mathbb{F}^k$ siempre que escribamos los vectores de $\mathbb{F}^k$ como columnas. Así , todo vector $v \in \mathbb{F}^k$ es una combinación lineal de las columnas de A, en otras palabras, v tiene la forma $v = Aw^T$ para algún vector $w \in \mathbb{F}^n$. Sea r la dimensión de $L^{\perp}$. Escojemos una base $b_1, ..., b_r$ de $L^{\perp}$ y a acompletamos a una base $b_1, ..., b_r, c_{r+1}, ..., c_n$ del espacio $\mathbb{F}^n$ (ver Proposición 4.21). Vamos a mostrar que los $n - r$ vectores

$$Ac_{r+1}^T, ..., Ac_n^T$$

forman una base del espacio $\mathbb{F}^k$. Probaremos que $k = n - r$. así, $r = n - k$, por lo cual la prueba estara terminada.

1. Los vectores de arriba generan $\mathbb{F}^k$. En efecto, todo vector $v \in \mathbb{F}^k$ tiene la forma

   $$v = Aw^T$$

   Podemos expresar el vector w como una combinación lineal

   $$w = \sum_{i=1}^{r} t_i b_i + \sum_{j=r+1}^{n} s_j c_j$$

   y entonces $Ab_i^T = 0$. implica

   $$v = A\left(\sum_{i=1}^{r} t_i b_i^T + \sum_{j=r+1}^{n} s_j c_j^T\right) = \sum_{j=r+1}^{n} s_j A_{c_j}^T$$

2. Los vectores de arriba son linealmente independientes. En efecto, consideremos una combinación lineal,

$$\sum_{j=r+1}^{n} t_j A c_j^T = 0^T$$

El vector $c = \sum_{j=r+1}^{n} t_j c_j$, cumple $Ac^T = 0^T$, es decir, $c \in L^{\perp}$. Así, c es una combinación lineal de los vetores $b_1, ..., b_r$ (forman una base de $L^{\perp}$) así como una combinación lineal de los vectores $c_{r+1}, ..., c_n$. Ya que $b_1, ..., b_r, c_{r+1}, ..., c_n$ forma una base de $\mathbb{F}^n$, la Proposición 4.20 implica que $c = 0$. Ahora, por la independencia lineal de $c_{r+1}, ..., c_n$, concluimos que $t_i = 0$ para todo i, lo cual prueba la independencia lineal de $Ac_{r+1}^T, ..., Ac_n^T$.

□

Capítulo 5

Anillo Cociente

En álgebra resulta de vital importancia la construcción del objeto cociente. En esta sección estudiaremos el anillo cociente, sus teoremas más clasicos y haremos énfasis en la retícula de ideales de los anillos cociente.

Proposición 5.1. *Sea R un anillo e I un ideal de R entonces $R/I = \{a+I \mid a \in R\}$ es un anillo con las siguientes operaciones:*

1. $(a + I) + (b + I) = (a + b) + I$.

2. $(a + I)(b + I) = (ab) + I$.

Además

$$\begin{aligned} \pi : R &\to R/I \\ a &\mapsto a + I \end{aligned}$$

es un momorfismo de anillos, llamado el momorfismo sobreyectivo natural.

Demostración. Veamos que la suma y la multiplicación están bien definidas, asumimos que $a+I = a'+I$ y $b+I = b'+I$ si y sólo si $a-a' \in I$ y $b-b' \in I$, así $ab-a'b' = ab-ba'+ba'-a'b' = b(a-a')+a'(b-b') \in I$ entonces tenemos que $ab - a'b' \in I$ si y sólo si $ab + I = a'b' + I$, así $(a + I)(b + I) = ab + I = a'b' + I = (a' + I)(b' + I)$. Veamos que $+$ es asociativa.
Sean $(a + I), (b + I), (c + I) \in R/I$, entonces $(a + I) + [(b + I) + (c + I)] = (a+I)+((b+c)+I) = [a+(b+c)]+I = [(a+b)+c]+I = [(a+b)+I]+(c+I) = [(a + I) + (b + I)] + (c + I)$.
Afirmamos que el elemento neutro es $(0 + I)$. Ya que

$$\begin{aligned}(0+I)+(a+I) &= (0+a)+I \\ &= a+I \\ &= (a+0)+I \\ &= (a+I)+(0+I)\end{aligned}$$

para toda $a \in R$.
Además $-(a+I) = -a+I$ para cada $a+I \in R/I$. Puesto que

$$\begin{aligned}(a+I)+(-a+I) &= (a+(-a))+I \\ &= (a-a)+I \\ &= 0+I\end{aligned}$$

Veamos la conmutatividad

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

Veamos la conmutatividad con el producto.

$$\begin{aligned}(a+I)(b+I) &= (ab)+I \\ &= (ba)+I \\ &= (b+I)(a+I)\end{aligned}$$

Veamos la asociatividad con el producto

$$\begin{aligned}[(a+I)(b+I)](c+I) &= [(ab)+I](c+I) \\ &= [(ab)c+I] \\ &= [a(bc)+I] \\ &= (a+I)[(bc)+I] \\ &= (a+I)[(b+I)(c+I)]\end{aligned}$$

Afirmamos que el elemento neutro para el producto es $1+I \in R/I$, pues

$$\begin{aligned}(a+I)(1+I) &= (a1)+I \\ &= (a1)+I \\ &= a+I \\ &= (1a)+I \\ &= (1+I)(a+I)\end{aligned}$$

Finalmente veamos la distribución,

$$\begin{aligned}(a+I)[(b+I)+(c+I)] &= (a+I)[(b+c)+I] \\ &= a(b+c)+I \\ &= (ab+ac)+I \\ &= (ab+I)+(ac+I) \\ &= (a+I)(b+I)+(a+I)(c+I).\end{aligned}$$

Veamos que $\pi : R \to R/I$ es momorfismo de anillos.

1. $\pi(a+b) = (a+b)+I = (a+I)+(b+I) = \pi(a)+\pi(b)$.
2. $\pi(ab) = (ab)+I = (a+I)(b+I) = \pi(a)\pi(b)$.
3. $\pi(1_R) = 1_R + I$.

□

Definición 5.2. *El anillo R/I construido en la Proposición 5.1 es llamado el* ***anillo cociente*** *de R módulo I.*

Definición 5.3. *Un* ***conjunto parcialmente ordenado (Copo)*** *es un par $(A, \leq)$ tal que*

1. *$a \leq a$ para todo $a \in A$.*
2. *Si $a \leq b$ y $b \leq c$, entonces $a \leq c$ para todo $a, b, c \in A$.*
3. *Si $a \leq b$ y $b \leq a$, entonces $a = b$ para todo $a, b \in A$.*

Si $(A, \leq)$ es un conjunto parcialmente ordenado y $B \subset A$ entonces:

1. *$a = infB$ si y sólo si para toda $b \in B$, $a \leq b$ y si $c \in A$ es tal que, para toda $b \in B$, $c \leq b$, entonces $c \leq a$.*

2. $a = supB$ *si y sólo si para toda* $b \in B$, $b \leq a$ *y si* $c \in A$ *es tal que, para toda* $b \in B$, $b \leq c$, *entonces* $a \leq c$.

Si $(A, \leq)$ *es un copo,* $(A, \leq)$ *se llama* ***retícula*** *si y sólo si para toda* $a, b \in A$ *existen* $a_0, b_0 \in A$ *tales que* $a_0 = inf\{a, b\}$ *y* $b_0 = sup\{a, b\}$.
Usaremos la siguiente notación $sup\{a, b\} = a \vee b$ *y* $inf\{a, b\} = a \wedge b$.

Definición 5.4. *Una función* $f : (A, \leq) \to (B, \leq)$ *entre dos copos es llamado un* ***morfismo de orden*** *si para dos elementos en* A, $a \leq a'$ *implica* $f(a) \leq f(a')$ *en* B. *Un morfismo de orden inyectivo se llamará morfismo de orden estricto. Una biyección que es un morfismo de orden, tal que su inversa es también un morfismo de orden, se llamara un* ***isomorfismo de orden****.*

Ejemplo 5.5. *Sea*

$$\begin{aligned}(\mathbb{N}, |) &\to (\mathbb{N}, \leq)\\ n \mid m &\mapsto n \leq m\end{aligned}$$

hay una correspondencia biyectiva de orden $(\to)$, *pero no es isomorfismo de ordenes parciales.*

Definición 5.6. *Una función* $f : L \to L'$ *entre dos retículas es llamado* ***morfismo de retículas*** *si:*

1. $f(a \vee b) = f(a) \vee f(b)$ *para todo* $a, b \in L$.

2. $f(a \wedge b) = f(a) \wedge f(b)$ *para todo* $a, b \in L$.

Un morfismo de retículas biyectivo es llamado un ***isomorfismo de retículas****. Para retículas isomorfas* L *y* L' *usamos la notación* $L \cong L'$.

Notación Sean R un anillo e I un ideal propio de R.

$[I, R] = \{K \subseteq R \mid I \leq K \leq R\}$.

$[\bar{0}, R/I] = \{L \subseteq R/I \mid L \leq R/I\}$.

Proposición 5.7. *Sea* R *un anillo y* $I \subseteq R$. *Entonces* $[I, R], [\bar{0}, R/I]$ *son retículas isomorfas.*

Demostración. Sea $\gamma : R \twoheadrightarrow R/I$ tal que $f(r) = r+I$. Definimos $\gamma' : [I, R] \to [\bar{0}, R/I]$ por $\gamma'(J) = \gamma(J)' = J/I$. Así que tenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccc} R & \xrightarrow{\gamma} & R/I \\ \uparrow & & \uparrow \\ J & \xrightarrow{\gamma'} & \gamma'(J) \end{array}$$

Veamos que γ' es inyectiva.
Sea $\gamma'(J) = \gamma'(K)$, entonces $\gamma(J) = \gamma(K)$, donde $\gamma(J), \gamma(K) \leq R/I$ con $J, K \in [I, R]$, entonces $\gamma^{-1}(\gamma(J)) = \gamma^{-1}(\gamma(K))$, luego $J + ker(\gamma) = K + ker(\gamma)$, es decir $J + I = K + I$, por lo tanto $J = K$.
Veamos que γ' es sobreyectiva.
Sea $\bar{J} \in [\bar{0}, R/I]$, así $\bar{0} \leq \bar{J}$, luego $ker(\gamma) = I = \gamma^{-1}(\bar{0}) \leq \gamma^{-1}(\bar{J})$, así $\gamma^{-1}(\bar{J}) \in [I, R]$ y $\gamma'(\gamma^{-1}(\bar{J})) = \gamma(\gamma^{-1}(\bar{J})) = \bar{J}$. Por lo tanto γ' es una biyección.
Finalmente debido a que γ' es inyectiva tenemos que:

1. $\gamma'(\bigcap_{j\in J} K_j) = \bigcap_{j\in J} \gamma'(K_j)$.
2. $\gamma'(\sum_{j\in J} K_j) = \sum_{j\in J} \gamma'(K_j)$.

Por lo tanto γ' es un isomorfismo de retículas.

□

Proposición 5.8. *Sea $f : R \to A$ un morfismo de anillos, entonces $kerf$ es un ideal de R, Imf es subanillo de A y*

$$R/kerf \cong Imf.$$

Demostración. Sea

$$\begin{aligned} \varphi : R/kerf &\to Imf \\ a + K &\mapsto f(a), \end{aligned}$$

donde $K = kerf$. Veamos que φ está bien definida.
Sea $a + K = b + K$ si y sólo si $a - b \in K$, es decir existe $k \in K$ tal que $a - b = k$ si y sólo si $a = b + k$. Por lo tanto

$$\varphi(a + K) = f(a) = f(b + k) = f(b) + f(k) = f(b) = \varphi(b + K).$$

Así φ está bien definida. Claramente φ es sobreyectiva. Veamos que φ es inyectiva.
Sea $\varphi(a+K) = \varphi(b+K)$ implica que $f(a) = f(b)$ si y sólo si $f(a) - f(b) = f(a-b) = 0$, así $a - b \in K$. Por lo tanto $a + K = b + K$. Ahora veamos que φ es momorfismo.

1. $\varphi(1_R + K) = f(1_R) = 1_A$.

2. $\varphi((a+K) + (b+K)) = \varphi((a+b)+K) = f(a+b) = f(a) + f(b) = \varphi(a+K) + \varphi(b+K)$.

3. $\varphi((a+K)(b+K)) = \varphi(ab+K) = f(ab) = f(a)f(b) = \varphi(a+K)\varphi(b+K)$.

□

Proposición 5.9. *Si $\langle 0\rangle \neq R$ un anillo, son equivalentes:*

1. *R es un campo*

2. *Sus únicos ideales son $\{0\}, \langle 1\rangle = R$*

3. *Todo morfismo distinto del morfismo cero de $R \to S$ es inyectivo.*

Demostración. 1. $\Rightarrow$]2. Sea $I \leq R$ con $I \neq \{0\}$ entonces existe $a \in I$ tal que $a \neq 0_R$, con a es unidad y $\langle a\rangle = \langle 1\rangle = R$.
2. $\Rightarrow$]3. Sea $f : R \to S$, con $Ker(f)$ un ideal propio de R, entonces $kerf = \{0\}$, por la Proposición 3.17, f es inyectivo.
3. $\Rightarrow$]1. Sea $0 \neq x \in R$, $\langle x\rangle$ un ideal propio de R tal que $\{\bar{0}\} \neq R/\langle x\rangle$, $\varphi : R \twoheadrightarrow R/\langle x\rangle$ el morfismo sobreyectivo natural. Entonces $\{0_R\} = Ker(\varphi) = \langle x\rangle$, luego $\langle x\rangle = \{0_R\}$, contradicción. Por lo tanto $\langle x\rangle = R$, luego x es unidad, notemos que $\varphi \neq 0$, por lo tanto R es campo. □

Proposición 5.10. *Sea $\mathbb{F}$ un campo, $p(x)$ un polinomio sobre $\mathbb{F}[x]$ con $\boldsymbol{gr}(p(x)) \geq 1$ y $\langle p(x)\rangle = I \leq \mathbb{F}[x]$. Entonces son equivalentes:*

1. *$p(x) \in \mathbb{F}[x]$ es irreducible.*

2. *$\mathbb{F}[x]/\langle f(x)\rangle$ es un campo.*

3. *$\mathbb{F}[x]/\langle f(x)\rangle$ es dominio entero.*

Demostración. Sea $I = \langle f(x) \rangle$.
$1 \Rightarrow 2$] Si $p(x)$ no es unidad, entonces $1 \notin I$ si y sólo si $0 + I \neq 1 + I$. Sea ahora $0 + I \neq f(x) + I \in \mathbb{F}[x]/I$, entonces $f(x) \notin I$ implica que $p(x) \nmid f(x)$ entonces $m.c.d(p(x), f(x)) = 1$, es decir, existen $s(x), t(x) \in \mathbb{F}[x]$ tal que $s(x)p(x) + t(x)f(x) = 1$, luego $[s(x)p(x) + t(x)f(x)] + I = 1 + I =$ $(s(x)+I)(p(x)+I)+(t(x)+I)(f(x)+I) = 1+I$ en $\mathbb{F}[x]/I$, pero $p(x)+I = 0$, así $(t(x)+I)(f(x)+I) = 1+I$. Por lo tanto $t(x)+I$ es el inverso de $f(x)+I$.
$2 \Rightarrow 3$] Por la proposición 1.16.
$3 \Rightarrow 1$] Supongamos que $p(x) \in \mathbb{F}[x]$ es no irreducible, es decir, existen $g(x), h(x) \in \mathbb{F}[x]$ tales que $p(x) = g(x)h(x)$ con $\mathbf{gr}(g), \mathbf{gr}(h) < \mathbf{gr}(p)$, así $g(x), h(x) \notin I$ entonces $g(x) + I \neq 0 + I \neq h(x) + I$, además $(g(x) + I)(h(x) + I) = g(x)h(x) + I = p(x) + I$ contradicción. Ya que $\mathbb{F}[x]/\langle f(x) \rangle$ es dominio entero y por lo tanto $p(x) \in \mathbb{F}[x]$ es irreducible. □

Proposición 5.11. *Sea k un campo, $p(x) \in k[x]$ mónico irreducible con $\mathbf{gr}(p(x) = d$, sea $I = \langle p(x) \rangle$, $K = k[x]/I$ y sea $\beta = x + I \in K$.*

1. *Entonces K es campo y $k' = \{a + I \mid a \in k\}$ es un subcampo de K isomorfo a k.*

2. *β es una raíz de $p(x)$ en K.*

Demostración. 1. El anillo cociente $K = k[x]/I$ es un campo, por la Proposición 5.10. Además tenemos el momorfismo sobreyectivo natural

$$\begin{aligned} \pi : k[x] &\to k[x]/I \\ f(x) &\mapsto f(x) + I \end{aligned}$$

Sea

$$\begin{aligned} \varphi = \pi|_k : k &\to k' \\ a &\mapsto a + I. \end{aligned}$$

Así $Im\varphi = k'$. Veamos que φ es inyectiva. Sea $\varphi(a) = \varphi(b)$ con $a, b \in k$, entonces $a + I = b + I$ si y sólo si $a - b \in I$. Supongamos que $a - b \neq 0$ entonces I contiene una unidad, así $I = k[x]$, contradicción, ya que $p(x)$ es mónico irreducible. Por lo tanto $a = b$. Por la Proposición 5.8, $k \cong k' \subseteq K$.

2. Sea $p(x) = a_0 + a_1x + ... + a_dx^{d-1} + x^d$, donde $a_i \in k$ para todo i. En $K = k[x]/I$, tenemos

$$\begin{aligned} p(\beta) &= (a_0 + I) + (a_1 + I)\beta + ... + (1 + I)\beta^d \\ &= (a_0 + I) + (a_1 + I)(x + I) + ... + (1 + I)(x + I)^d \\ &= (a + I) + ... + (a_1x + I) + ... + (1x^d + I) \\ &= a_0 + a_1x + ... + x^d + I \\ &= p(x) + I = I, \end{aligned}$$

ya que $p(x) \in I = \langle p(x) \rangle$. Pero $I = 0 + I$ es el elemento cero de $K = k[x]/I$, y así β en k es una raíz de $p(x)$.

□

Definición 5.12. *Sea K un campo y $k \subseteq K$ un subcampo. Entonces K es una extensión de campo sobre k y lo denotamos por K/k. Notemos que K/k es un espacio vectorial sobre k. Decimos que K/k es una extensión finita si $dim_k K < \infty$.*

Notación: $[K : k] = dim(K/k) < \infty$ es el grado de la extensión.

Proposición 5.13. *Sea k un campo y $f(x) \in k[x]$ con $\boldsymbol{gr}(f(x)) \geq 1$, entonces existe K/k una extensión de campo tal que $f(x) \in K[x]$ es producto de polinomios lineales.*

Demostración. La demostración se hará por inducción sobre $\mathbf{gr}(f(x)) = n \geq 1$. Si $n = 1$, entonces $K = k$. Si $n > 1$, escribimos $f(x) = p(x)g(x)$, donde $p(x)$ es mónico irreducible. Ahora por la Proposición 5.11 proporciona un campo F que contiene a k tal que $p(x)$ tiene una raíz $a \in \mathbb{F}$, entonces $p(x) = (x - a)h(x) \in \mathbb{F}[x]$. Por lo tanto, en $F[x]$, tenemos $p(x) = (x - a)h(x)$ y $f(x) = (x-a)h(x)g(x)$. Sea $\mathbf{gr}(h(x)g(x)) = n-1$, por hipótesis de inducción existe $F \subseteq K($ y por lo tanto $k \subset K)$ y así $h(x)g(x) \in K$ es producto de factores lineales. Por lo tanto $f(x) = (x - a)h(x)g(x) \in K[x]$ es producto de factores lineales. □

Capítulo 6

Ideales Primos y Máximos

En esta sección trataremos con la generalización de número primo, dando lugar a los conceptos de ideal primo e ideal máximo. Introducimos conceptos muy importantes en la teoría de anillos tales como: el espectro, el radical y el nilradical. Además de mostrar la existencia de ideales máximos.

Definición 6.1. *Un ideal propio, P de R se dice que es ideal primo, si para cualesquiera $a, b \in R$ tal que $ab \in P$ y $a \notin P$, implica que $b \in P$.*

Ejemplo 6.2. *Recordemos que un anillo R distinto de cero es un dominio entero si y sólo si $ab = 0$ en R, implica que $a = 0$ ó $b = 0$. Por lo tanto, el ideal $(0) = \{0\}$ es primo en R si y sólo si R es un dominio entero.*

Ejemplo 6.3. *Un ideal $(m) = m\mathbb{Z} < \mathbb{Z}$ es un ideal primo si y sólo si m es un primo o cero.*

Demostración. $\Rightarrow$] Dado que m y $-m$ generan el mismo ideal principal, consideremos solo los generadores positivos. Si $m = 0$ entonces el resultado se sigue del ejemplo anterior. Ya que $\mathbb{Z}$ es un dominio entero. Si $m > 0$, veamos que (m) es un ideal propio, de otra manera $1 \in (m)$, así, debe existir un entero p tal que $mp = 1$, contradicción. Luego, si $ab \in (m)$ entonces $p \mid ab$. Por el Lema de Euclides, $p \mid a$ ó $p \mid b$, es decir, $p \in (a)$ ó $p \in (b)$. Por lo tanto (p) es un ideal primo.
$\Leftarrow$] Si $m > 1$ y no es primo, entonces se tiene una factorización $m = ab$ con $0 < a, b < m$. Por lo tanto ni a ni b es un múltiplo de m, es decir $a \notin (m)$ y $b \notin (m)$ pero $ab = m \in (m)$ y así (m) no es un ideal primo. □

Definición 6.4. *Un ideal propio M en R es llamado un ideal máximo si no existe un ideal propio de R, digamos J, tal que*

$$M < J < R$$

Ejemplo 6.5. *Si $R = \mathbb{F}$ es campo, entonces su único ideal máximo es* (0).

Proposición 6.6. *Sea R un anillo. El ideal $M < R$ es máximo si y sólo si R/M es un campo.*

Demostración. $\Rightarrow$] Sea $a + M \neq 0 + M$, lo que significa que $a \notin M$, ahora consideremos el ideal $Ra + M$ tal que $M < Ra + M$, por ser M máximo, entonces $M + Ra = R$, luego existen $m \in M$, $b \in R$ tal que $1 = m + ab$, entonces $ab - 1 \in M$, así $ab + M = 1 + M$, de manera que $(a+M)^{-1} = b+M$, luego R/M es un campo.
$\Leftarrow$] Sea $M < I$ con I un ideal de R, luego existe $a \in I$ tal que $a \notin M$, por ser R/M campo, entonces existe $b \in M$ talque $(a + M)(b + M) = 1 + M$, luego $1 - ab \in M < I$ y como $ab \in I$ entonces $1 \in I$ e $I = R$. □

Ejemplo 6.7. *Una aplicación del teorema anterior es la siguiente:*

$$p\mathbb{Z} = (p) \text{ es máximo si y sólo si } \mathbb{Z}_p \text{ es campo.}$$

Proposición 6.8. *Sea R un anillo. Un ideal propio P es un ideal primo si y sólo si R/P es un dominio entero.*

Demostración. $\Rightarrow$. Supongamos que $(a + P)(b + P) = 0 + P$, luego $ab \in P$, por ser P primo. $a \in P$ ó $b \in P$, es decir, $a + P = 0 + P$ ó $b + P = 0 + P$.
$\Leftarrow$. Supongamos que $ab \in P$, entonces $(a + P)(b + P) = 0 + P$. Por ser R/P dominio entero, $a + P = 0 + P$ ó $b + P = 0 + P$, es decir, $a \in P$ ó $b \in P$. □

Proposición 6.9. *Sea R un anillo. Si M es un ideal máximo, entonces es un ideal primo.*

Demostración. Por la Proposición 1.16. □

Ejemplo 6.10. *En general no es cierto el recíproco del teorema anterior. Por ejemplo, en $\mathbb{Z}$, $\{0\}$ es ideal primo pero no es ideal máximo.*

Proposición 6.11. *Si $0 \neq R$ es un anillo. Entonces R tiene un ideal máximo. Más aún, todo ideal propio I en R está contenido en un ideal máximo.*

Demostración. Nótese que la segunda afirmación implica la primera. Sea $I < R$ y X la familia de todos los ideales propios que contienen a I. Notemos que $X \neq \emptyset$ ya que $I \in X$. Es claro que X está ordenado parcialmente y el orden parcial en X está dado por la inclusión. Sea $\mathcal{C}$ una cadena en X. Es decir, dados $I, J \in \mathcal{C}$, entonces $I \subseteq J$ ó $J \subseteq I$. Afirmamos que

$$I^* = \bigcup_{I \in \mathcal{C}} I$$

es una cota superior de $\mathcal{C}$. Claramente, $I \subseteq I^*$, así que resta demostrar que $I^* \in X$, es decir, que I^* es un ideal propio.

1. I^* es un ideal. Supongamos que $r \in R$ y $a \in I^*$, entonces, $a \in I$ para algún $I \in \mathcal{C}$. Por lo tanto, $ra \in I$, y así $ra \in I^*$.
   Ahora supongamos que $a, b \in I^*$, entonces existen $I_a, I_b \in \mathcal{C}$, con $a \in I_a$ y $b \in I_b$, pero ya que $\mathcal{C}$ es totalmente ordenado, $I_a \subseteq I_b$ ó $I_b \subseteq I_a$. Sin perdida de generalidad, supongamos que $I_b \subseteq I_a$, entonces $a, b \in I_a$. Por lo tanto $a + b \in I_a$, así $a + b \in I^*$.
   Claramente $0 \in I$ para todo $I \in \mathcal{C}$. Por lo tanto $0 \in I^*$.

2. I^* es propio, es decir, $I^* \neq R$. Ya que $1 \notin J$, para cada $J \in \mathcal{C}$, tenemos que $1 \notin I^*$, de ahí que $I^* \neq R$.
   Ya que X satisface las hipótesis del Lema de Zorn, concluimos que tiene elemento máximo, el cual es un ideal propio máximo que contiene a I.

□

Definición 6.12. *El conjunto de todos los ideales primos en un anillo R es llamado el* ***espectro*** *de R y lo denotaremos por $Spec(R)$ mientras el conjunto de sus ideales máximos es el* ***espectro máximo*** *de R, denotado por $Specm(R)$. Además se puede observar que $Specm(R) \subseteq Spec(R)$.*

Ejemplo 6.13. *1. Si $R = \mathbb{Z}$. Entonces* ***Spec***$(\mathbb{Z}) = \{(p) \mid p \in \mathbb{Z}$ *primo*$\} \cup \{(0)\}$

Definición 6.14. *Un anillo R con un único ideal máximo M es llamado* ***anillo local****. Por la Proposición 6.6*

$$\mathcal{K} := R/M$$

es un campo. A este campo se le conoce como el ***campo residual****. Se verificará más adelante que $M = \{r \in R \mid r$ no es una unidad$\}$.*
Un momorfismo de anillos locales $f : R \to S$ es llamado un morfismo local si $f(M_R) \subseteq M_S$, donde M_R y M_S son los ideales máximos de los anillos locales de R y S respectivamente.

Ejemplo 6.15. *Un morfismo local es el siguiente:*

$$\mu : \mathbb{Z}_4 \to \mathbb{Z}_2$$

$$0 \mapsto 0$$

$$1 \mapsto 1$$

$$2 \mapsto 0$$

$$3 \mapsto 1$$

Este morfismo está determinado por la estructura de $\mathbb{Z}_4$, es decir, $\mathbb{Z}_4$ es un anillo finito, conmutativo, local con ideal máximo $2\mathbb{Z}_4 = \{0, 2\}$ y campo residual $\mathbb{Z}_4/2\mathbb{Z}_4 \cong \mathbb{Z}_2$. Por lo tanto, μ es el morfismo sobreyectivo natural y el ideal máximo $2\mathbb{Z}_4$ asigna el elemento cero del campo.

Proposición 6.16. *Sea R un anillo.*

1. *Si $M \neq (0)$ un ideal tal que cada $x \in R \setminus M$ es una unidad. Entonces R es un anillo local y M es su único ideal máximo.*

2. *Si M es un ideal máximo tal que cada elemento del conjunto $1 + M := \{1 + x \mid x \in M\}$ es una unidad en R. Entonces R es un anillo local.*

Demostración. 1. Sea I un ideal propio. Supongamos que existe $x \in I$ y $x \notin M$, luego $x \in R \setminus M$, entonces x es unidad. Así que $I = R$, contradicción. Por lo tanto $x \in I$ implica $x \in M$. Así $I \subseteq M$. De esta manera demostramos que M es el único ideal máximo.

2. Sea $x \in R \setminus M$. Como M es máximo, el ideal $J = (x, M)$ coincide con R, esto implica que existe $y \in R$ y $t \in M$ tal que $xy + t = 1$, así $xy = 1 - t \in 1 + M$ es una unidad en R, luego existe $s \in R$ tal que $(xy)s = s(xy) = 1$. Así $1 = x(ys) = (ys)x$. Por lo tanto $x \in U(R)$. Por 1. R es un anillo local. □

Definición 6.17. *Un anillo que contiene solo un número finito de ideales máximos es llamado* ***semilocal****.*

Ejemplo 6.18. *1. Cada anillo finito es un anillo semilocal.*

2. $\mathbb{Z}$ tiene infinitos ideales máximos y por lo tanto no es un anillo semilocal.

3. $\mathbb{Q}$ es infinito y es semilocal.

Definición 6.19. *Dado un ideal I de R, el conjunto*

$$\sqrt{I} := \{r \in R \mid r^s \in I \text{ para algún } s \in \mathbb{N}\}$$

es un ideal de R, llamado el ***radical*** *de I. Claramete $I \subseteq \sqrt{I}$. Un ideal J es llamado un ideal radical si $\sqrt{J} = J$.*

Ejemplo 6.20. *El ideal radical del ideal $4\mathbb{Z}$ es $2\mathbb{Z}$.*

Proposición 6.21. *Sea M un ideal máximo de R. Entonces M es un ideal radical.*

Demostración. Por definición, $M \subseteq \sqrt{M}$. Por ser M máximo se tiene que $M = \sqrt{M}$ en este caso M es el ideal radical ó $\sqrt{M} = R$, contradicción. □

Proposición 6.22. *Si P es un ideal primo de R, entonces P es un ideal radical.*

Demostración. Por definición $P \subseteq \sqrt{P}$. Sea a un elemento de $\sqrt{P}$ y n un entero positivo tal que $a^n \in P$. Demostraremos que si $a^n \in P$, entonces $a \in P$, se hará por inducción.

1. Para $n = 1$, $x = x^1 \in P$ implica que $x \in P$.
2. Supongamos que se cumple para $n = k$, es decir, que $a^k \in P$ implica que $a \in P$.
3. Veamos que se cumple para $n = k+1$, $a^{k+1} = aa^k$ están en el ideal primo, entonces $a \in P$ ó $a^k \in P$, si $a \in P$, se termina la demostración. Si $a^k \in P$, entonces por hipótesis de inducción, tenemos $a \in P$.

□

Definición 6.23. *Si R es un anillo conmutativo, entonces su* ***nilradical*** *$nil(R)$ es definido como la intersección de todos los ideales primos en R, es decir*

$$Nil(R) := \bigcap_{P \in Spec(R)} P$$

Definición 6.24. *Si R es un anillo, entonces el* ***radical de Jacobson*** *$J(R)$ está definido como la intersección de todos los ideales máximos de R.*

Ejemplo 6.25. *1. En cualquier dominio entero el nilradical es el ideal (0).*

2. En cualquier anillo local el radical de Jacobson es el único ideal máximo.

Proposición 6.26. *Sea R un anillo, entonces*

$$Nil(R) = \{f \in R \mid \exists n \in \mathbb{N}, f^n = 0\}.$$

Demostración. Supongamos que S es el conjunto de los elementos nilpotentes. Veamos $S \subseteq Nil(R)$. Sea $f \in S$ y $p \in spec(R)$, entonces existe $n \in \mathbb{N}$ tal que $f^n = 0$. Pero $0 \in P$ así que $f^n \in P$ por la Proposición 6.22. Así que $f \in P$ para todo $p \in Spec(R)$, entonces $f \in \bigcap_{p \in Spec(R)} P = Nil(R)$. Así $S \subseteq Nil(R)$. Ahora veamos $Nil(R) \subseteq S$. Es decir, debemos demostrar que $f \in Nil(R)$ implica que existe $n \in \mathbb{N}$ tal que $f^n = 0$. Pero notemos que esto es equivalente a demostrar que si $f^n \neq 0$, para $n \in \mathbb{N}$ entonces existe $p \in spec(R)$ tal que $f \notin P$.
Supongamos que $f^n \neq 0$ para todo $n \in \mathbb{N}$. Sea $\Gamma = \{I \leq R \mid f^n \notin I,$ para algún $n \in \mathbb{N}\}$. Note que $\Gamma \neq \emptyset$, ya que si $\{0\} \subseteq R$ y $n \in \mathbb{N}$, entonces $f^n \notin \{0\}$, así $\{0\} \in \Gamma$. Además $(\Gamma, \subseteq)$ es un copo. Consideremos una cadena $\Phi \subseteq \Gamma$, $\bigcup \Phi \subseteq R$ la unión de una cadena de ideales es un ideal, pero la unión de ideales no necesariamos es ideal. Supongamos que existe $n \in \mathbb{N}$ tal que $f^n \in \bigcup \Phi$, entonces $f^n \in I$ para algún $I \in \Phi$. Contradicción. Así $\bigcup \Phi \in \Gamma$ y es una cota superior para Φ. Por el Lema de Zorn, Γ tiene un elemento máximo P. Veamos que P es ideal primo.
Supongamos que $x, y \notin P$. Veamos que $xy \notin P$. Como $x \notin P$ entonces $P < P + \langle x \rangle$ es ideal mayor propio. De manera análoga $y \notin P$, entonces $P < P + \langle y \rangle$. Como P es máximo, entonces $P + \langle x \rangle \notin \Gamma$ y $P + \langle y \rangle \notin \Gamma$, entonces, existen $n, m \in \mathbb{N}$ tales que $f^n \in P + \langle x \rangle$ y $f^m \in P + \langle y \rangle$, así $f^n = p_1 + r_1 x$ y $f^m = p_2 + r_2 y$, con $r_1, r_2 \in R$. Así

$$f^n f^m = (p_1 + r_1 x)(p_2 + r_2 y) = p_1 p_2 + p_1 r_2 y + p_2 r_1 x + r_1 x r_2 y.$$

Entonces $f^n f^m = f^{n+m} \in P + \langle xy \rangle$. Así $P + \langle xy \rangle \notin \Gamma$ y $P + \langle xy \rangle \geq P$. Así $P + \langle xy \rangle > P$ entonces $xy \notin P$. Por lo tanto P es un ideal primo, $P \in \Gamma$ y $f^n \notin P$ para todo $n \in \mathbb{N}$. En particular, para $n = 1$, $f \notin P$ entonces $f \notin Nil(R)$. Por lo tanto $Nil(R) \subseteq S$. Así $Nil(R) = S$.

□

Proposición 6.27. *Sea R un anillo, $x \in J(R)$ si y sólo si $1 - xy$ es una unidad de R para cada $y \in R$.*

Demostración. $\Rightarrow$] Si $1 - xy$ no es una unidad, entonces por la Proposición 6.11, $(1-xy)$ pertenece a algún ideal máximo M de R; ya que $x \in J(R) \subseteq M$, $xy \in M$, lo que implica $1 \in M$, contradicción.
$\Leftarrow$] Si $x \notin M$, para algún ideal máximo M, entonces $(M, x) = R$ por ser M máximo. Por lo tanto, existe $v \in M$ y $y \in R$ tal que $v + xy = 1$. Así $1 - xy \in M$, lo que implica que $1 - xy$ no es una unidad, contradicción. $\square$

Capítulo 7

Estructura de los campos finitos

Dado que la teoria de códigos se desarrolla sobre los campos finitos, en esta sección se estudiaran algunas de sus características pero todavia más importante se demostrará la abundante existencia de dichos campos.

Proposición 7.1. *Para todo elemento β de un campo finito $\mathbb{F}$ con q elementos, $\beta^q = \beta$.*

Demostración. En el caso en que $\beta = 0$ se tiene $0^q = 0$. Supongamos ahora que $\beta \neq 0$. Sea $\mathbb{F}^* = \{\beta_1, ..., \beta_{q-1}\}$ el conjunto de todos los elementos no cero de $\mathbb{F}$. Dado que $\beta \neq 0$, tenemos que $\{\beta\beta_1, \beta\beta_2, ..., \beta\beta_{q-1}\} = \mathbb{F}^*$ son también elementos no cero, luego $\beta_1\beta_2...\beta_{q-1} = \beta^{q-1}(\beta_1\beta_2...\beta_{q-1})$. Por lo tanto $\beta^{q-1} = 1$.

□

Proposición 7.2. *Sea F un subcampo de E con $|F| = q$. Entonces un elemento β de E está en F si y sólo si $\beta^q = \beta$.*

Demostración. $\Rightarrow$] Por la Proposición 7.1 se cumple.
$\Leftarrow$] Consideremos el polinomio $x^q - x \in E[x]$. El cual tiene a lo más q raíces distintas en E. Como todos los elementos de F son raíces de $x^q - x$ y $|F| = q$, obtenemos $F = \{\alpha \mid \alpha \text{ es raíz de } x^q - x \text{ en } E\}$. Por lo tanto, para todo $\beta \in E$ que satisface $\beta^q = \beta$, es una raíz de $x^q - x$, es decir, $\beta \in F$. □

Proposición 7.3. *Sea $p \in \mathbb{Z}$ primo, y $n \in \mathbb{N}$. Entonces existen campos con p^n elementos.*

Demostración. Sea $q = p^n$, y consideremos $f(x) = x^q - x \in \mathbb{F}_p[x]$. Por la Proposición 5.13 existe K campo tal que $\mathbb{F}_p \subseteq K$ en el cual $f(x)$ se descompone totalmente.
Este polinomio tiene q raíces distintas en K ya que su derivada es $qx^{q-1}-1 = -1$ en $\mathbb{F}_q[x]$ y así no puede tener una raíz común con x^q-x por la Proposición 2.25. Sea $S = \{a \in K \mid a^q - a = 0\}$. Entonces S es un subanillo de K ya que

1. $f(1) = 1^q - 1 = 0$, así $1 \in S$.
2. Si $a, b \in S$ entonces por la Proposición 1.25 que $(a-b)^q = a^q - b^q = a - b$ y así $a - b \in S$.
3. Para $a, b \in S$ y $b \neq 0$, tenemos $(ab^{-1})^q = a^q(b^{-1})^q = a^q b^{-q} = a^q(b^q)^{-1} = ab^{-1}$ y así $ab^{-1} \in S$.

Pero, por otro lado $x^q - x$ se puede factorizar en S, ya que S contiene toda sus raíces. Por lo tanto $K = S$, y ya que S tiene q elementos, K es campo finito con $q = p^n$ elementos. □

Definición 7.4. *Un elemento α en un campo finito $\mathbb{F}_q$ es llamado **elemento generador** o **primitivo** de $\mathbb{F}_q$ si $\mathbb{F}_q = \{0, \alpha, \alpha^2, ..., \alpha^{q-1}\}$.*

Ejemplo 7.5. *En $\mathbb{Z}_7$ el elemento $\bar{3}$ es un elemento primitivo.*

Definición 7.6. *El **orden** de un elemento distinto de cero $\alpha \in \mathbb{F}_q$ denotado por **ord**(α) es el entero positivo más pequeño k tal que $\alpha^k = 1$.*

Ejemplo 7.7. *Siguiendo con el Ejemplo 7.5, el orden de $\bar{3}$ es 6.*

Proposición 7.8. *El orden **ord**(α) divide a $q - 1$ para todo $\alpha \in \mathbb{F}_q^*$.*

Demostración. Sea m un entero positivo que tenga como propiedad $\alpha^m = 1$, usando el algoritmo de la división escribimos $m = a\mathbf{ord}(\alpha) + b$ para algunos enteros a y b tales que $0 \leq a$ y $0 \leq b < \mathbf{ord}(\alpha)$ entonces

$$1 = \alpha^m = \alpha^{a\mathbf{ord}(\alpha)+b} = \alpha^{a\mathbf{ord}(\alpha)}\alpha^b = (\alpha^{\mathbf{ord}})^a\alpha^b = 1\alpha^b = \alpha^b.$$

Por lo tanto $b = 0$, esto es, $a\mathbf{ord}(\alpha) = m$, es decir, $\mathbf{ord}(\alpha) \mid m$. Además sabemos que para cada $\alpha \in \mathbb{F}_q^*$ tenemos que $\alpha^{q-1} = 1$, es decir, $\mathbf{ord}(\alpha) \mid q - 1$. □

Proposición 7.9. *Para dos elementos $\alpha, \beta \in \mathbb{F}_q^*$ si m.c.d(ord(α),ord(β)) = 1, entonces ord($\alpha\beta$)=ord(α)ord(β).*

Demostración. Sea $r = \text{ord}(\alpha)\text{ord}(\beta)$. Es claro que $\alpha^r = 1 = \beta^r$ ya que el $\text{ord}(\alpha)$ y $\text{ord}(\beta)$ son divisores de r. Por lo tanto

$$(\alpha\beta)^r = \alpha^r\beta^r = 1.$$

Así $\text{ord}(\alpha\beta) \leq \text{ord}(\alpha)\text{ord}(\beta)$.
Por otro lado, sea $t = \text{ord}(\alpha\beta)$, tenemos

$$1 = (\alpha\beta)^{t\cdot ord(\alpha)} = (\alpha^{ord(\alpha)})^t\beta^{t\cdot ord(\alpha)} = \beta^{t\cdot ord(\alpha)}.$$

Esto implica que $\text{ord}(\beta)$ divide $t\cdot\text{ord}(\alpha)$ por la Proposición de 7.8, así el $\text{ord}(\beta)$ divide a t ya que $\text{ord}(\alpha)$ es primo relativo con $\text{ord}(\beta)$.
En la misma forma se muestra que $\text{ord}(\alpha)$ divide a t. Esto implica que $\text{ord}(\alpha)$ y $\text{ord}(\beta)$ dividen a t. Por lo tanto, $\text{ord}(\alpha\beta) = t \geq \text{ord}(\alpha)\text{ord}(\beta)$. □

Proposición 7.10. *Un elemento distinto de cero de $\mathbb{F}_q$ es un elemento primitivo si y sólo si su orden es $q-1$.*

Demostración. Por la Proposición 7.1 tenemos que para todo $\alpha \in \mathbb{F}_q^*$ se tiene que el orden es $q-1$ si y sólo si $\alpha, \alpha^2, ..., \alpha^{q-1}$ son distintos. Esto es equivalente a decir que $\mathbb{F}_q = \{0, \alpha, \alpha^2, ..., \alpha^{q-1}\}$. □

Proposición 7.11. *Todo campo finito tiene al menos un elemento primitivo.*

Demostración. Sea m el mínimo común múltiplo de los ordenes de todos los elementos de $\mathbb{F}_q^*$. Si r^k es una potencia de un primo en la factorización canónica de m, entonces $r^k \mid \text{ord}(\alpha)$ para algún $\alpha \in \mathbb{F}_q^*$. El orden de $\alpha^{ord(\alpha)/r^k}$ es r^k. Por lo tanto, si

$$m = r_1^{k_1}...r_n^{k_n}$$

es la factorización canónica de m para distintos primos $r_1, ..., r_n$, entonces, para cada $i = 1, ..., n$, existen $\beta_i \in \mathbb{F}_q^*$ con $\text{ord}(\beta_i) = r_i^{k_i}$. La Proposición 7.9 implica que existe $\beta \in \mathbb{F}_q^*$ con $\text{ord}(\beta) = m$. Además $m \mid (q-1)$ por la Proposición 7.8 y por otra parte todos los $(q-1)$ elementos de $\mathbb{F}_q^*$ son raíces del polinomio $x^m - 1$, de modo que $m \geq q-1$. Por lo tanto, $\text{ord}(\beta) = m = q-1$ y el resultado se termina por la Proposición 7.10. □

Capítulo 8

Codificación

En esta sección, empezaremos a definir desde el punto matemático conceptos tales como código, mensaje, codificación y decodificación que son de vital importancia para la teoría de códigos. Para esto, usamos como punto de partida las definiciones dadas por J. Adámek ([1]) que sirvierón como inspiración para que nosotros pudieramos dar desde nuestro punto de vista nuestras propias definiciones.

Definición 8.1. *Dados conjuntos finitos A (alfabeto de origen) y B (alfabeto código) una palabra origen es un elemento de $\bigcup_{n\in\mathbb{N}} A^n$, una palabra código es un elemento de $\bigcup_{n\in\mathbb{N}} B^n$.*
Al conjunto de las palabras origen lo denotamos por P_o y al conjunto de palabras códigos P_c.

Definición 8.2. *Si $p \in P_o$ es una palabra origen, entonces $p \in A^m$ para algún $m \in \mathbb{N}$. Es decir, $p = (a_1, ..., a_m)$, con $a_i \in A$, en este caso decimos que p tiene longitud m.*
Una codificación es una función inyectiva $f : A \to P_c$ la cual asigna a cada elemento de A exactamente una palabra código. Si $f : A \to P_c$ es una codificación, decimos que $f(A)$ es un código.

Definición 8.3. *Sea f una codificación. Sea $a \in P_o$, entonces $a = (a_1, ..., a_n)$ para algún $n \in N$, un mensaje de la codificación f enviado por a es un elemento m de P_c definido por $m = f'(a) = (f(a_1), ..., f(a_n))$, donde $f' : P_o \to P_c$.*

Definición 8.4. *La decodificación del mensaje m es*

$$d_m = (f^{-1}f(a_1), ..., f^{-1}f(a_n)).$$

Ejemplo 8.5. *Sea $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ y $B = \mathbb{F}_2$, y sea P_c de las palabras código de longitud 5 que tengan solo dos 1's, de tal manera que la columna de la izquierda representa a los elementos de A y la columna de la derecha representa a su imagen bajo la codificación f representada por esta asignación.*

$$\begin{gathered} 1 \mapsto (1,1,0,0,0) \\ 2 \mapsto (1,0,1,0,0) \\ 3 \mapsto (0,1,1,0,0) \\ 4 \mapsto (1,0,0,1,0) \\ 5 \mapsto (0,1,0,1,0) \\ 6 \mapsto (0,0,1,1,0) \\ 7 \mapsto (1,0,0,0,1) \\ 8 \mapsto (0,1,0,0,1) \\ 9 \mapsto (0,0,1,0,1) \\ 10 \mapsto (0,0,0,1,1) \end{gathered}$$

así, la palabra $a = (1, 7, 3)$ al ser enviada mediante f', obtenemos el mensaje $m = f'(a) = ((1,1,0,0,0),(1,0,0,0,1),(0,1,1,0,0))$. Usualmente se abusa de la notación y el mensaje m se denonta $m = (11000, 10001, 01100)$. Esto se hace para evitar una notación engorrosa.

A veces en la teoría de códigos se omite la codificación y se escribe simplemente el código $C = \{c_1, ..., c_n\}$ donde $c_i \in B^n$. Por ejemplo, podemos escribir el código del ejemplo anterior como $C = \{c_1 = (1,1,0,0,0), c_2 = (1,0,1,0,0), ..., c_{10} = (0,0,0,1,1)\}$. No se escribe la codificación debido a que es bastante claro que podemos definir cualquier función codificadora. En el ejemplo, escribimos $A = \{1, ..., 10\}$, donde f es tal que $f(i) = c_i$; pero podríamos haber codificado también así: $A = \{a, b, c, d, e, f, g, h, i, j\}$ donde f es tal que $f(a) = c_1$, $f(b) = c_2$,...,$f(j) = c_{10}$. Notemos que las dos funciones nos propocionan exactamente el mismo código.

Definición 8.6. *Si B es un alfabeto código definimos al conjunto de palabras código de longitud n como B^n. Si $f : A \to B^n$ es una codificación, decimos que $f(A)$ es un código bloque de longitud n.*
Además, si $B = \mathbb{F}_2$ y $f : A \to \mathbb{F}_2^n$ es una codificación, entonces diremos que $f(A)$ es un código binario de longitud n.

Capítulo 9

Detección de errores

Así, para decodificar será útil poner una medida de que tan cerca dos palabras código están una de la otra. Una manera de hacerlo es con la distancía de Hamming que a continuación definimos. Además definiremos otro concepto desde el punto matemático, el concepto de transmisión y veremos cuando un código se puede corregir y cuando no.

Definición 9.1. *Sean x e y dos palabras de longitud n sobre el mismo alfabeto $\mathcal{B}$. La* ***distancia de Hamming*** *entre x e y, denotada por $d(x, y)$, se define como el número de coordenadas en que x e y difieren, es decir, $d : \mathcal{B}^n \times \mathcal{B}^n \longrightarrow [0, n] \subset \mathbb{N}$. Donde si $x = (x_1, x_2, ..., x_n)$ e $y = (y_1, y_2, ..., y_n)$, entonces*

$$d(x, y) = \boldsymbol{d}(x_1, y_1) + \boldsymbol{d}(x_2, y_2) + ... + \boldsymbol{d}(x_n, y_n)$$

con

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

A partir de aquí, si C es un código de longitud n. Escribiremos $x = x_1x_2...x_n \in C$ en lugar de $x = (x_1, x_2, ..., x_n)$.

Ejemplo 9.2. *Sea $\mathbb{Z}_2 = \{0, 1\}$ y sea $x = 01010$, $y = 01101$, $z = 11101 \in \mathbb{Z}_2^5$. Entonces*

1. $d(x, y) = 3$
2. $d(y, z) = 1$
3. $d(z, x) = 4$

Ahora veamos que $(\mathcal{B}^n, d)$ es un espacio métrico.

Proposición 9.3. *Sean x, y, z, palabras de longitud n sobre $\mathcal{B}$. Entonces tenemos que:*

I. $0 \leq d(x, y) \leq n$

II. $d(x, y) = 0$ *si y sólo si* $x = y$.

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

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

Demostración. Consideremos $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ y $z = (z_1, z_2, ..., z_n)$.

I. De la definición de distancia de Hamming tenemos que $d(x, y) = d(x_1, y_1) + d(x_2, y_2) + ... + d(x_n, y_n)$, supongamos que en todos los casos $x_i = y_i$ por lo que $d(x, y) = 0$ y ahora suponiendo que $x_i \neq y_i$, para todo $i \in \{1, 2, ..., n\}$, tendríamos que $d(x_1, y_1) + d(x_2, y_2) + ... + d(x_n, y_n) = 1 + 1 + ... + 1 = 1(n) = n$ por lo que $0 \leq d(x, y) \leq n$.

II. $d(x, y) = d(x_1, y_1) + d(x_2, y_2) + ... + d(x_n, y_n) = 0$ si y sólo si $d(x_i, y_i) = 0$ para todo $i \in \{1, ..., n\}$ si y sólo si $x_i = y_i$ para todo $i \in \{1, ..., n\}$ si y sólo si $x = y$.

III. $d(x, y) = d(x_1, y_1) + d(x_2, y_2) + ... + d(x_n, y_n) = d(y_1, x_1) + d(y_2, x_2) + ... + d(y_n, x_n) = d(y, x)$.

IV. Si $x = z$ entonces $d(x, z) = 0$. Observemos que si x y z difieren en un lugar, es decir, $x_i \neq z_i$ para algún $i \in \{i, ...n\}$ entonces x y y difieren en ese mismo lugar $x_i \neq y_i$ ó z e y difieren en ese mismo lugar $z_i \neq y_i$ ó en ambos, de ahí el número de lugares donde x y z que difiere es menor o igual a el número de lugares donde x e y difieren más el número de lugares donde z e y difieren.

□

Apartir de aquí, cuando mencionemos distancia nos estaremos refiriendo a la distancia de Hamming.

Definición 9.4. *Sea $C \subseteq \mathcal{B}^n$ un código bloque de tamaño n, definimos una* ***transmisión*** *como una función $T : C \longrightarrow \mathcal{B}^n$, además decimos que T no tiene ruido si $T(c) = c$ para todo $c \in C$, es decir, si T es la función inclusión, de otra manera decimos que T tiene ruido. Decimos que $c \in C$ tiene k errores en la transmisión T si $d(c, T(c)) = k$.*

La idea básica de detección de errores, es simple, si recibimos una palabra que no es una palabra código entonces sabremos que se ha cometido un error.

Definición 9.5. *Dado un código bloque C de tamaño n definimos la* ***distancia mínima*** *de C denotada $d(C)$, como la menor distancia no nula entre cualesquiera dos palabras código, es decir,*

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

Definición 9.6. *Dado $x \in \mathcal{B}^n$, con $\mid \mathcal{B} \mid = q$ y $r \geq 0$, se define la* ***esfera*** *de radio r y centrada en x como*

$$S_q(x, r) = \{y \in \mathcal{B}^n : d(x, y) = r\}.$$

y la bola de radio r centrada en x como

$$B_q(x, r) = \{y \in \mathcal{B}^n : d(x, y) \leq r\} = \bigcup_{i=0}^{r} S_q(x, i)$$

Veamos que las bolas de radio $t = \left\lfloor \frac{d(C)-1}{2} \right\rfloor$ centradas en palabras código son disjuntas. Donde $\lfloor u \rfloor$ es el mayor entero menor que o igual a u.

Proposición 9.7. *Si C es un código con distancia mínima $d(C) = 2t + 1$ ó $d(C) = 2t + 2$, entonces*

$$B_q(c, t) \cap B_q(c', t) = \emptyset$$

Para todo $c, c' \in C$ con $c \neq c'$.

Demostración. Sea $x \in B_q(c, t)$ con $c \in C$. Entonces $x \notin B_q(c', t)$ para todo $c' \in C$ con $c' \neq c$. Ya que si no fuera así, por la desigualdad del triángulo tendriamos

$$d(c, c') \leq d(c, x) + d(x, c') \leq t + t = 2t < 2t + 1 = d(C).$$

Lo que es una contradicción. □

Definición 9.8. *Un código bloque C se dice que* ***detecta*** *t* ***errores*** *siempre que para cada palabra código a y cada palabra a' obtenida de a por diferir en $1, 2, ..., t$ simbolos, a' no es una palabra código.*

Proposición 9.9. *Un código C detecta t errores si y sólo si $d(C) \geq t + 1$.*

Demostración. $\Rightarrow$] Supongamos que C detecta t errores y que $d(C) < t+1$, es decir, $d(C) \leq t$, entonces existen $x, y \in C$ tales que $1 \leq d(x, y) = d(C) \leq t$. Por lo tanto es posible que comencemos con la palabra código x y $d(C)$ errores de tal manera que la palabra código resultante es y una palabra código en C. Por lo tanto C, no detecta t errores.
$\Leftarrow$] Supongamos que $d(C) \geq t + 1$. Si $c \in C$ y x son tales que $1 \leq d(c, x) \leq t < d(C)$. Se sigue que $x \notin C$. Por lo tanto C detecta t errores. □

Definición 9.10. *Un código de bloque C se dice que* ***corrige*** *t* ***errores*** *siempre que para cada palabra código a y cada palabra a' obtenido por diferir en $1, 2, ..., t$ símbolos, la distancia de Hamming $d(a, a')$ es estrictamente más pequeña que la distancia de Hamming de a' a cualquier otra palabra código. En símbolos, para cada palabra código $a \in C$ y cada palabra a' tal que $1 \leq d(a, a') \leq t$, se tiene que $d(a, a') < d(b, a')$ para todas las palabras código $b \in C$, $b \neq a$.*

Proposición 9.11. *Un código C corrige t errores si y sólo si $d(C) \geq 2t+1$.*

Demostración. $\Rightarrow$]Sea $d(C) \leq 2t$. Entonces mostraremos que C no puede corregir t errores distribuidos arbitrariamente. En otras palabras, encontraremos una palabra código a y una palabra a' que tiene distancía de Hamming t o menor de a y sin embargo, su distancía de Hamming de una palabra código diferente es aún más pequeña. Sean $a, b \in C$ con $d(a, b) = d(C)$. Sean $i_1, i_2..., i_r$ todos los índices en los cuales a difiere de b. Entonces $r \leq 2t$. Supongamos que enviamos a y el ruido cambia todos los símbolos a_i con $i \in \{i_2, i_4, i_6, ...\}$ (es decir, todos los índices pares en los cuales a y b difieren) a los valores en b. Esto es, recibimos la palabra a', donde

$$a_i{}' = \begin{cases} a_i = b_i & si \quad i \neq i_1, i_2, ..., i_r, \\ a_i & si \quad i = i_1, i_3, ..., \\ b_i & si \quad i = i_2, i_4, ... \end{cases}$$

Así $d(a, a') \leq \frac{r}{2} \leq t$ y sin embargo $d(a', b) \leq d(a', a)$. Esto puede llevar a la decodificación de a' incorrectamente como b.

$\Leftarrow]$ Sea $d(C) > 2t$. Entonces C corrige t errores. En efecto, supongamos que enviamos una palabra a y recibimos una palabra a' con distancia de Hamming $d(a, a') \leq t$. Tenemos la siguiente situación con cualquier palabra código $b \neq a : d(a, b) \geq d(C) > 2t$ y por la desigualdad del triángulo

$$d(a, a') + d(a', b) \geq d(a, b) > 2t. \tag{9.1}$$

Por lo tanto $d(a', b) > 2t - d(a, a') \geq 2t - t = t \geq d(a, a')$. □

Ejemplo 9.12. *El código* $C_1 = \{00, 01, 10, 11\}$ *no detecta ningún error, mientras que* $C_2 = \{001, 010, 100\}$ *detecta 1 error,*

Notación Un código de longitud n, con M palabras códigos, y con distancia mínima $d = d(C)$, es llamado un **(n,M,d) código.**

Proposición 9.13. *Sea* C *un código. Una bola* $B_q(x, c)$ *de radio* r *centrada en una palabra código* c *tiene:*

$$\binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + ... + \binom{n}{r}(q-1)^r$$

elementos.

Demostración. Primero calculemos el número de vectores que tienen una distancia 1 de la palabra código c. Estos vectores son los que difieren en C en exactamente una coordenada, además hay n posibles coordenadas y $q - 1$ formas para hacer una entrada diferente. Por lo tanto el número de vectores que tienen una distancia de 1 a c es $n(q - 1)$. Ahora calculemos el número de vectores que tienen distancia m de c , hay $\binom{n}{m}$ formas en las cuales podemos elegir m coordenadas para diferir de los valores de c, además para cada una de esas coordenadas, hay $q - 1$ elecciones para símbolos diferentes de el correspondiente símbolo de la palabra código c. Por lo tanto, hay

$$\binom{n}{m}(q-1)^m$$

elementos que tienen distancia m de la palabra código c. Incluyendo la misma palabra código c y usando la identidad $\binom{n}{0} = 1$. Obtenemos:

$$\binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + ... + \binom{n}{r}(q-1)^r \ .$$

□

Proposición 9.14. *Sea* C *un* ***[n,M,d]****-código con* $d \geq 2t + 1$ *Entonces*

$$M \leq \frac{q^n}{\sum_{j=0}^{t} \binom{n}{j}(q-1)^j}$$

Demostración. Alrededor de cada palabra código colocamos una bola de radio t. Ya que la mínima distancia de el código es $d \geq 2t + 1$ las bolas no se superponen. El número total de vectores de cada bola no puede ser más grande que q^n. Por lo tanto obtenemos

$$\text{(número de palabras códigos)} \times \text{(número de elementos por bola)}$$

$$= M \sum_{j=0}^{t} \binom{n}{j} (q-1)^j \leq q^n$$

Por lo tanto se cumple la desigualdad deseada. □

A este resultado se le llama la **Cota de Hamming** o a veces también llamado **empaquetamiento de esferas.**

Capítulo 10

Códigos Lineales

En está sección hablaremos sobre qué son los códigos lineales y discutiremos sus propiedades elementales.

Definición 10.1. *Un* ***código lineal*** *de longitud n sobre un campo $\mathbb{F}_q$ es un subespacio $C \subset \mathbb{F}_q^n$ de dimensión k. En este caso decimos que C es un* ***[n,k]****-código sobre $\mathbb{F}_q$. Si C tiene distancía mínima d entonces decimos que C es un* ***[n,k,d]****-código sobre $\mathbb{F}_q$.*

Observación 10.2. *Por lo tanto un código lineal de longitud n es un conjunto de palabras de longitud n tal que:*

1. *Si a y b son palabras códigos, entonces $a + b$ es una palabra código.*
2. *Si a es una palabra código, entonces para cada múltiplo escalar t, ta es una palabra código.*
3. *Si C es un subespacio de dimensión k de $\mathbb{F}_q^n$ y sean $e_1, ..., e_k$ una base de C. Entonces para cada palabra código v se sigue que,*

$$v = \sum_{n=1}^{k} u_i e_i.$$

4. *Si el alfabeto código F tiene r símbolos, entonces existen r^k palabras códigos en C.*

Ejemplo 10.3. *Sea $\mathbb{Z}_2^3$. $C = \{000, 001, 010, 011\}$ es un código lineal.*

Definición 10.4. *Sea C un código lineal en $\mathbb{F}_q^n$.*

1. *El **código dual** de C, denotado por $C^\perp$, es el complemento ortogonal del subespacio C de $\mathbb{F}_q^n$.*
2. *La **dimensión** del código lineal C es la dimensión de C como un espacio vectorial sobre $\mathbb{F}_q$, denotada por dim(C).*

Proposición 10.5. *Sea C un código lineal de longitud n sobre $\mathbb{F}_q$. Entonces,*

1. $|C| = q^{dim(C)}$, *es decir,* $dim(C) = \boldsymbol{log}_q |C|$.
2. $C^\perp$ *es un código lineal y* $dim(C) + dim(C^\perp) = n$.
3. $(C^\perp)^\perp = C$.

Demostración. 1. Por el Lema 4.26 se cumple.

2. Usando la igualdad del Lema 4.26 y la Proposición 4.27 con $C = S$.

3. Veamos que $C \subset (C^\perp)^\perp$. Sea $c \in C$, debemos demostrar que $c \in (C^\perp)^\perp$, es decir $c \cdot x = 0$ para todo $x \in C^\perp$. Así tomando $c \in C$ y $x \in C^\perp$ se tiene que $x \cdot c' = 0$ Para todo $c' \in C$, en particular para c. Por lo tanto $x \cdot c = 0$. Luego usando la igualdad de la parte 2. y reemplazando C por $C^\perp$. tenemos que la $dim(C^\perp)^\perp = n - (n - k) = k = dim(C)$. Por lo tanto $C = (C^\perp)^\perp$.

□

Ejemplo 10.6. *Sea $\mathbb{Z}_2^4$. Consideremos el código $C = \{0000, 1010, 0101, 1111\}$, así, $dim(C) = \boldsymbol{log}_2 |C| = \boldsymbol{log}_2 4 = 2$. Es claro que $C^\perp = \{0000, 0101, 1010, 1111\} = C$, es decir, $dim(C^\perp) = 2$. Además se cumplen las propiedades 2. y 3. de la Proposición 10.5.*

Definición 10.7. *Sea $\boldsymbol{x} \in \mathbb{F}_q^n$ una palabra. El **peso de Hamming** de $\boldsymbol{x}$, denotado por $w(\boldsymbol{x})$, está definido como el número de coordenadas no nulas de $\boldsymbol{x}$, es decir,*

$$w(\boldsymbol{x}) = d(\boldsymbol{x}, \boldsymbol{0}),$$

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

Lema 10.8. *Para todo elemento x de $\mathbb{F}_q$, podemos definir el peso de Hamming como sigue:*

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

es decir el peso de $\boldsymbol{x}$ *es la distancia de* $\boldsymbol{x}$ *al* $\boldsymbol{0} = (0, 0, ..., 0)$. *Si* C *es un código se define el* ***peso*** *de* C *como,*

$$w(C) = min\{w(\boldsymbol{x}) : \boldsymbol{x} \neq 0, \boldsymbol{x} \in C\}.$$

El peso Hamming de $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ *puede reescribirse como,*

$$w(\boldsymbol{x}) = w(x_1) + w(x_2) + ... + w(x_n)$$

Proposición 10.9. *Si* x,$y \in \mathbb{F}_q^n$, *entonces* $d(x, y) = w(x - y)$

Demostración. Para x, $y \in \mathbb{F}_q$, $d(x, y) = 0$ si y sólo si $x = y$, pero esto es equivalente a $x - y = 0$ por otro lado $w(x - y) = d(x - y, 0) = 0$. Por lo tanto $d(x, y) = w(x - y)$.

□

Proposición 10.10. *Sea* q *par. Si* $x, y \in \mathbb{F}_q^n$, *entonces* $d(x, y) = w(x + y)$.

Demostración. Ya que $a = -a$ para toda $a \in \mathbb{F}_q$ cuando q es par, el resultado es una inmediata consecuencia de la Proposición 10.9. □

Definición 10.11. *Si* $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_3) \in \mathbb{F}_2^n$ *palabras binarias se define la intersección de* x *e* y *como*

$$x \cap y = (x_1y_1, x_2y_2, ..., x_ny_n)$$

Lema 10.12. *Si* $x, y \in \mathbb{F}_2^n$ *entonces* $w(x + y) = w(x) + w(y) - 2w(x \cap y)$.

Demostración. Es suficiente ver los siguientes 4 casos, para las coordenadas de x y y.

x	y	$w(x) + w(y) - 2w(x \cap y)$	$w(x + y)$
0	0	0	0
0	1	1	1
1	0	1	1
1	1	0	0

□

Lema 10.13. *Para cualquier potencia prima* q *y* $x, y \in \mathbb{F}_q^n$ *tenemos que*

$$w(x) + w(y) \geq w(x + y) \geq w(x) - w(y)$$

Demostración. Sea $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ donde

$$w(x) = d(x,0) = d(x_1,0) + d(x_2,0) + ... + d(x_n,0)$$
$$w(y) = d(y,0) = d(y_1,0) + d(y_2,0) + ... + d(y_n,0)$$

entonces

$w(x)+w(y) = d(x_1,0)+d(x_2,0)+...+d(x_n,0)+d(y_1,0)+d(y_2,0)+....+d(y_n,0)$

pero $d(x_i,y_i) \leq d(x_i,0) + d(0,y_i)$. Así

$w(x+y) = d(x,y) = \sum_{i=1}^{n} d(x_i,y_i) \leq \sum_{i=1}^{n} d(x_i,0) + d(0,y_i) = w(x) + w(y)$.

Análogamente

$$\begin{aligned} w(x) - w(y) &= d(x_1,0) + ... + d(x_n,0) - d(y_1,0) - ... - d(y_n,0) \\ &= (d(x_1,0) - d(0,y_1)) + ... + (d(x_n,0) - d(y_n,0)) \end{aligned}$$

Recordando que $w(x+y) = d(x,y) = d(x_1,y_1) + ... + d(x_n,y_n)$, debemos demostrar que $d(x_i,y_i) \geq d(x_i,0) - d(y_i,0)$ para toda $i \in \{1,...,n\}$.
Supongamos que $d(x_i,y_i) < d(x_i,0) - d(y_i,0)$, luego $d(x_i,y_i) + d(y_i,0) < d(x_i,0)$, pero $d(x_i,0) \leq d(x_i,y_i) + d(y_i,0) < d(x_i,0)$, lo que es una contradicción. Así $d(x_i,y_i) \geq d(x_i,0) - d(y_i,0)$. Por lo tanto

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

□

Proposición 10.14. *Si C es un código lineal entonces $d(C) = w(C)$.*

Demostración. Usando la definición de distancia mínima existen $x, y, z \in C$ con $x \neq y$ y $z \neq 0$ tales que,

$$d(C) = d(x,y) = w(x-y) \geqslant w(C)$$

puesto que $x - y$ es un elemento de C, por otro lado

$$w(C) = w(z) = d(z,0) \geqslant d(C)$$

□

Ejemplo 10.15. *Sea $\mathbb{Z}_2^4$ y el código lineal binario $C = \{0000, 1000, 0100, 1100\}$.*

$$w(1000) = 1.$$
$$w(0100) = 1.$$
$$w(1100) = 2.$$

Por lo tanto, d(C) = 1.

En teoria de códigos una base para un código lineal es a menudo representado en la forma de una matriz, llamada una **matriz generadora**, mientras una matriz que representa una base para el código dual es llamada una **matriz checadora de paridad.**

Definición 10.16. *Sea C un **[n,k,d]**-código con $c_1, c_2, ..., c_k$ una base. Una **matriz generadora** de C es una matriz $G \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ definida por*

$$G = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix}$$

cuyas filas forman una base para C.

Observación 10.17. *Notemos que G siempre existe y tiene rango k. Observemos que G genera a C, es decir,*

$$C = \{uG : u \in \mathbb{F}_q^k\}$$

Sea $\{c_1, ..., c_k\}$ una base de C y $c_i = \sum_{j=1}^{n} c_{ij}e_j$ para ciertos $c_{ij} \in \mathbb{F}_q$, con $e_1, ...e_n$ la base canónica de $\mathbb{F}_q^n$. Si $u \in \mathbb{F}_q^k$, entonces $uG \in C$, pues

$$uG = (u_1, u_2, ..., u_k) \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix} =$$

$$= (\sum_{i=1}^{k} u_i c_{i1},\ ...,\ \sum_{i=1}^{k} u_i c_{in}) = \sum_{j=1}^{n} (\sum_{i=1}^{k} u_i c_{ij}) e_j = \sum_{i=1}^{k} u_i (\sum_{j=1}^{n} c_{ij} e_j) = \sum_{i=1}^{k} u_i c_i.$$

Recíprocamente, si $c \in C$*, existen únicos* $u_1, ..., u_k$ *tal que* $c = \sum_{i=1}^{k} u_i c_i$*. Luego, tomando* $u = (u_1, ..., u_k) \in \mathbb{F}_q^k$ *se cumple que* $c = uG$*.*
Por lo tanto, el proceso de representación de los elementos $u \in \mathbb{F}_q^k$ *como palabras código* $c = uG$ *en* C *es llamado* ***codificación****.*

Ejemplo 10.18. *Sea* C *un* ***[5,3]****-código lineal binario con matriz generadora*

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

Como las filas son linealmente independientes, C *tiene rango 3. Además*

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} x_1, & x_2, & x_1 + x_3, & x_1 + x_2, & x_2 + x_3 \end{pmatrix}.$$

Entonces el mensaje $u = 101$ *es codificado como 10011.*

Definición 10.19. *Sea* C *un* ***[n,k]****-código. Una matriz* H *sobre* $\mathbb{F}_q$ *se dice* ***matriz checadora de paridad*** *de* C *siempre que se cumpla lo siguiente: Para palabras código* $v = v_1 v_2 ... v_n$ *en* $\mathbb{F}^n$*,* v *es una palabra código de* C *si y sólo si* $Hv^T = 0^T$*, es decir,*

$$v \in C \Leftrightarrow H \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ . \\ . \\ . \\ v_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

Definición 10.20.

1. *Una matriz generadora está en* ***forma estándar*** *si es de la forma* $G = (I_k|A)$ *donde* I_k *es la matriz identidad* $k \times k$ *y* A *es* $k \times (n-k)$*.*

2. *Una matriz checadora de paridad está en* ***forma estándar*** *si es de la forma* $(Y|I_{n-k})$ *donde* I_{n-k} *es la matriz identidad* $(n-k) \times (n-k)$*.*

Lema 10.21. *Sea C un **[n,k]**-código lineal sobre $\mathbb{F}_q$ con matriz generadora G. Entonces $v \in \mathbb{F}_q^n$ pertence a $C^\perp$ si y sólo si $\boldsymbol{v}$ es ortogonal para toda fila de G. Es decir $\boldsymbol{v} \in C^\perp$ si y sólo si $vG^T = 0$. En particular, dado una $(n-k) \times n$ matriz H, entonces H es una matriz checadora de paridad para C si y sólo si las filas de H son linealmente independientes y $HG^T = 0$. Donde G^T representa la transpuesta de G.*

Demostración. Sea r_i la i-ésima fila de G. En particular $r_i \in C$ para todo $1 \leq i \leq k$, y todo $c \in C$ puede ser escrito como

$$c = \lambda_1 r_1 + ... + \lambda_k r_k.$$

donde $\lambda_1, ..., \lambda_k \in \mathbb{F}_q$.
$\Rightarrow$] Si $\mathbf{v} \in C^\perp$, entonces $\mathbf{v} \cdot c = 0$ para todo $c \in C$. En particular, $\mathbf{v}$ es ortogonal a r_i para todo $1 \leq i \leq k$, es decir $\mathbf{v}G^T = \mathbf{0}$.
$\Leftarrow$] Si $\mathbf{v} \cdot r_i = 0$ para todo $1 \leq i \leq k$, entonces para cualquier $c = \lambda_1 r_1 + ... + \lambda_k r_k \in C$,

$$\mathbf{v} \cdot c = \lambda_1 (v \cdot r_1) + ... + \lambda_k (v \cdot r_k) = 0$$

Para la última afirmación, si H es la matriz checadora de paridad para C, entonces las filas de H son linealmentes independientes por definición. Ya que las filas de H son palabras código en $C^\perp$, se deduce de la afirmación anterior que $HG^T = 0$.
Reciprocamente, si $HG^T = 0$, entonces la afirmación anterior muestra que las filas de H, y por lo tanto el espacio fila de H, están contenidos en $C^\perp$. Ya que las filas de H son linealmente independientes, el espacio fila de H tiene dimensión $n-k$, entonces el espacio fila de H es de hecho $C^\perp$. En otras palabras, H es una matriz checadora de paridad para C. □

Observación 10.22. *Una formulación alternativa pero equivalente para el Lema 10.21 es la siguiente:*
*Sea C un **[n,k]**-código lineal sobre $\mathbb{F}_q$, con matriz checadora de paridad H. Entonces $\boldsymbol{v} \in \mathbb{F}_q^n$ pertenece a C si y sólo si $\boldsymbol{v}$ es ortogonal a toda fila de H; es decir, $\boldsymbol{v} \in C$ si y sólo si $\boldsymbol{v}H^T = \boldsymbol{0}$. En particular, dado una $k \times n$ matriz G, entonces G es una matriz generadora para C si y sólo si las filas de G son linealmente independientes y $GH^T = 0$.*

Proposición 10.23. *Si $G = (I_k \mid B)$ es la matriz generadora en forma estándar de un **[n,k]**-código C, entonces una matriz checadora de paridad para C es $H = (-B^T \mid I_{n-k})$*

Demostración. Sea C un (n,k)-código con matriz generadora $G = (I_k \mid B)$. Si L denota el espacio fila de la matriz H de arriba, entonces vamos a mostrar que C es el complemento ortogonal de L. Primero observemos que el rango de la matriz H es $n-k$, ya que I_{n-k} es una matriz identidad. Por consiguiente, la dimensión de L^T es $n-(n-k)=k$ es igual que la dimensión de C. Por lo tanto es suficiente mostrar que toda fila g de la matriz G cumple $Hg^T = 0^T$, luego se sigue que toda palabra código v de C también cumple $Hv^T = 0^T$, es decir, que C es un subespacio de $L^\perp$. Entonces la igualdad de las dimensiones implica que $C = L^\perp$. Ahora, queremos probar que HG^T es la matriz 0. Así

$$GH^T = [-B^T \mid I'] \begin{bmatrix} I \\ - \\ B^T \end{bmatrix} = -B^T I + I' B^T = -B^T + B^T = 0,$$

donde $I' = I_{n-k}$. □

10.1. Equivalencia de códigos lineales

A continuación vamos a dar la definición tradicional de equivalencia de códigos.

Definición 10.24. *Dos **[n,M]**-códigos C_1 y C_2 sobre $\mathbb{F}_q$ son **equivalentes**, y se denotan por $C_1 \simeq C_2$, si existe una permutación $\sigma \in \mathbb{S}_n$ de las n coordenadas y permutaciones $\pi_1, ..., \pi_n \in Biy(\mathcal{B})$ del alfabeto, tales que*

$$c_1, c_2, ..., c_n \in C_1 \text{ si y sólo si } \pi_1(c_{\sigma(1)})\pi_2(c_{\sigma(2)})...\pi_n(c_{\sigma(n)}) \in C_2.$$

Una definición alternativa es la siguiente.

Definición 10.25. *Dos **[n,M]**-códigos sobre $\mathbb{F}_q$ son **equivalentes** si un código puede ser obtenido de otro por una serie de las siguientes operaciones:*

1. *Permutaciones en las coordenadas de los códigos.*
2. *Multiplicación por algún escalar distinto de cero en alguna coordenada del código.*

Ejemplo 10.26. *Sea $\mathbb{Z}_3^3$ y consideremos el código $C = \{000, 011, 022\}$. Permutando la primera y la segunda coordenada, y seguida por la multiplicación por 2 en la tercera coodenada, obtenemos el código*

$$C' = \{000, 102, 201\}$$

Proposición 10.27. *Todo código lineal es equivalente a un código en forma estándar.*

Demostración. La matriz generadora G de un (n, k)-código lineal C tiene rango k, por lo tanto tiene k columnas que son linealmente independientes. Supongamos que las primera k columnas de G son linealmente independientes, esto es $G = [A \mid B]$, donde A es una matriz cuadrada de $k \times k$. Luego aplicando operaciones elementales sobre las filas, las cuales transforman a A en la matriz identidad. Luego si aplicamos las misma operaciones elementales sobre las filas de G, obtenemos una matriz $G' = [I \mid B']$. Ya que G' es también una matriz generadora de el código C, se sigue que C está en forma estándar. □

10.2. Decodificación de códigos lineales.

A partir de ahora veremos dos formas de tratar de recuperar la palabra código, una vez que ya ha sido enviado por el canar de transmición. Estamos hablando de la decodificación del vecino más cercano y la decodificación por Síndrome.

Definición 10.28. *Sea C un código lineal de longitud n sobre $\mathbb{F}_q$ y sea $u \in \mathbb{F}_q^n$ cualquier vector de longitud n, definimos la* ***clase de*** *C determinada por u como el conjunto*

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

Ejemplo 10.29. *Sea $\mathbb{Z}_2^3$ y $C = \{000, 101, 010, 111\}$. Entonces*

$$\begin{aligned}
C + 000 &= \{000, 101, 010, 111\},\\
C + 001 &= \{001, 100, 011, 110\},\\
C + 010 &= \{010, 111, 000, 101\},\\
C + 011 &= \{011, 110, 001, 100\},\\
C + 100 &= \{100, 001, 110, 011\},\\
C + 101 &= \{101, 000, 111, 010\},\\
C + 110 &= \{110, 011, 100, 001\},\\
C + 111 &= \{111, 010, 101, 000\}.
\end{aligned}$$

Notemos que

$$C + 000 = C + 010 = C + 101 = C + 111 = C;$$
$$C + 001 = C + 011 = C + 100 = C + 110 = \mathbb{Z}_2^3/C.$$

Proposición 10.30. *Sea C un **[n,k,d]**-código lineal sobre un campo finito $\mathbb{F}_q$. Entonces*

1. *Todo vector de $\mathbb{F}_q^n$ está contenido en alguna clase de C.*
2. *Si $|C| = r$, entonces $|C + v| = r$.*
3. *Para todo $u, v \in \mathbb{F}_q^n$, $u \in C + v$ implica que $C + u = C + v$.*
4. *Dos clases son la misma o su intersección es vacía.*
5. *Para todo $u, v \in \mathbb{F}_q^n$, $u - v \in C$ si y sólo si u y v están en la misma clase.*

Demostración. 1. Cada elemento $x \in \mathbb{F}_q^n$ está en la clase $C + x$. En efecto, C contiene al $\mathbf{0}$, y $x = \mathbf{0} + x$.

2. Sea $C = \{c_1, c_2, ..., c_r\}$ con $c_i \neq c_j$ para $i \neq j$. Toda clase tiene la forma:

$$C + x = \{c_1 + x, c_2 + x, ..., c_r + x\},$$

y $c_i + x \neq c_j + x$, para $i \neq j$. En efecto, si $c_i + x = c_j + x$, entonces restándo x obtenemos $c_i = c_j$. Por lo tanto, $C + x$ tiene r elementos.

3. Sea $x \in C + u$ entonces $x = c + u$ para $c \in C$, como $u \in C + v$ entonces $u = c' + v$ para $c' \in C$. Luego $x = (c + c') + v$ está en $C + v$, ya que $c + c' \in C$. Entonces por 2. $C + u = C + v$.

4. Si las clases $C + x$ y $C + y$ tienen un elemento en común digamos z, entonces escribimos z en dos formas

$$z = c' + x = c'' + y \text{ para } c', c'' \in C.$$

Ahora, demostraremos que todo elemento t de la clase $C + x$ está en $C + y$; por simetría llegamos a la conclusión de que $C + x = C + y$. Expresamos t como $t = c + x$ para algún $c \in C$. Entonces

$$t = c + x = c + (c'' - c' + y) = (c + c'' - c') + y.$$

Como C es un código lineal $(c + c'' - c') \in C$. Por lo tanto $t \in C + y$.

5. $\Rightarrow$] Si $u - v \in C$, entonces existe $c \in C$ tal que $c = u - v$, en otras palabras $u = c + v \in C + v$. Concluimos que u y v están en la misma clase $v + C$. $\Leftarrow$] Si u y v están en la misma clase, digamos $C + w$, entonces $u = c + w$ para algún $c \in C$ y $v = c' + w$ para algún $c' \in C$. Entonces C contiene $c - c'$, y tenemos que

$$u - v = c + w - (c' + w) = c - c' \in C.$$

□

Definición 10.31. *Una palabra de mínimo peso de Hamming en una clase es llamado* ***lider de clase****.*

Definición 10.32. *Sea $\mathbb{F}_q^n$. Una* ***matriz estándar*** *para un* ***[n,k]****-código es una matriz de $(q^{n-k} \times q^k)$ donde,*

1. *La primera fila contiene todas las palabras código.*
2. *Cada fila tiene una clase con el lider de clase en la primera columna.*
3. *La entrada en la i-ésima fila y la j-ésima columna es la suma del i-ésimo lider de clase y la j-ésima palabra código.*

Ejemplo 10.33. *Las clases para el código $C = \{0000, 1011, 0101, 1110\}$ son los siguientes:*

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

La cual forma una matriz estándar.

10.3. Decodificación del vecino más cercano para códigos lineales.

Sea C un código lineal. Asumamos que la palabra código $\mathbf{v}$ es transmitida y la palabra $\mathbf{w}$ es recibida, resultando en el **patrón de error**.

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

Entonces $\mathbf{w}-\mathbf{e}=\mathbf{v}\in C$, así por la parte 5. de la Proposición 10.30, el patrón de error e y la palabra recibida w están en la misma clase. **La decodificación del vecino más cernano** trabaja de la siguiente manera: Al recibir la palabra w, escogemos una palabra e de menor peso en la clase $w+C$ y concluimos que $v=w-e$ fue la palabra transmitida.

Observación 10.34. *Interpretamos $e+C=\{e+v \mid v\in C\}$ como el conjunto de todas las posibles palabras recibidas $w=e+v$, cuando una palabra código v es enviada y el canal ruidoso agrega el patrón de error e.*

Ejemplo 10.35. *Sea $\mathbb{Z}_2^4$ y $C=\{0000,1011,0101,1110\}$. Decodificamos la siguiente palabra recibida: $\boldsymbol{w}$=1101.*
Primero, escribimos la matriz estándar de C, que es exactamente la misma que en el Ejemplo 10.33:

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

$\boldsymbol{w}$=1101: $w+C$ está en la cuarta clase. La palabra de menor peso en esta clase es 1000*. Por lo tanto* $1101-1000=1101+1000=0101$ *fue la palabra código que más probabilidad tuvo de ser enviada.*

10.4. Decodificación por Síndrome

Definición 10.36. *Sea C un **[n,k,d]**-código lineal sobre $\mathbb{F}_q$ y sea H una matriz checadora de paridad para C. Para cualquier $\boldsymbol{w}\in\mathbb{F}_q^n$, el **síndrome** de $\boldsymbol{w}$ es la palabra $S(\boldsymbol{w})=\boldsymbol{w}H^T$.*

Proposición 10.37. *Sea C un **[n,k,d]**-código lineal y sea H una matriz checadora de paridad para C. Para $\boldsymbol{u},\boldsymbol{v}\in\mathbb{F}_q^n$ tenemos:*

I. $S(\boldsymbol{u}+\boldsymbol{v})=S(\boldsymbol{u})+S(\boldsymbol{v})$

II. $S(\boldsymbol{u})=\boldsymbol{0}$ *si y sólo si* $\boldsymbol{u}$ *es una palabra código en C.*

III. $S(\boldsymbol{u})=S(\boldsymbol{v})$ *si y sólo si* $\boldsymbol{u}$ *y* $\boldsymbol{v}$ *están en la misma clase se C.*

Demostración. I. $S(\mathbf{u}+\mathbf{v})=(\mathbf{u}+\mathbf{v})H^T=\mathbf{u}H^T+\mathbf{v}H^T=S(\mathbf{u})+S(\mathbf{v})$.

II. Por definición de síndrome, $S(\mathbf{u}) = \mathbf{0}$ si y sólo si $\mathbf{u}H^T = \mathbf{0}$, por la Observación 10.22, tenemos que $\mathbf{u} \in C$.

III. $S(\mathbf{u}) = S(\mathbf{v})$ si y sólo si $S(\mathbf{u}) - S(\mathbf{v}) = S(\mathbf{u} - \mathbf{v}) = \mathbf{0}$, así $\mathbf{u} - \mathbf{v} \in C$. Por la propiedad 5. de la Proposición 10.30 $\mathbf{u}$ y $\mathbf{v}$ están en la misma clase.

□

Proposición 10.38. *El patron de error tiene el mismo síndrome como la palabra recibida. Esto es, para cada palabra código $\boldsymbol{v}$ y cada patrón de error $\boldsymbol{e}$ la palabra $\boldsymbol{w} = \boldsymbol{v} + \boldsymbol{e}$ cumple*

$$\boldsymbol{e}H^T = \boldsymbol{w}H^T \tag{10.1}$$

Demostración. Ya que la palabra código cumple $\mathbf{v}H^T = \mathbf{0}$ tenemos,

$$\mathbf{w}H^T = (\mathbf{v} + \mathbf{e})H^T = \mathbf{v}H^T + \mathbf{e}H^T = \mathbf{0} + \mathbf{e}H^t = \mathbf{e}H^T.$$

□

Esto último significa que todas las palabras en la clase tienen el mismo síndrome.

10.5. Códigos Cíclicos

En está sección vamos estudiar de manera resumida algunas de las propiedades de los códigos cíclicos que son un caso particular de los códigos lineales. El cual tienen mucho intéres ya que se pueden estudiar con ayuda de los anillos de polinomios.
Recordemos que si X es un conjunto con n elementos entonces $S_n = \{\sigma : X \to X \mid \sigma \text{ es función biyectiva}\}$ es el grupo simétrico de orden $n!$. Recordemos $\sigma = (1\ 2\ 3\ ...\ n)$ es ciclo de longitud n.

Definición 10.39. *Sea C un código lineal, $c = (a_1, a_2, ...a_n) \in C$ y $\sigma = (1\ 2\ ...\ n) \in S_n$, luego definimos*

$$\begin{aligned}\sigma(c) &= (a_{\sigma(1)}, a_{\sigma(2)}, ..., a_{\sigma(n)}) \\ &= (a_2, a_3, ..., a_n, a_1).\end{aligned}$$

Decimos que C es un código ciclico si $\sigma(c) \in C$, para toda $c \in C$.

Supongamos el código C sobre un campo $\mathbb{F}_q$. Sea el correspondiente elemento $v_0, v_1, ..., v_{n-1}$ en C con el elemento $v(x) = v_0 + v_1x + ... + v_{n-1}x^{n-1}$ en el anillo cociente $R_n = \mathbb{F}_q[x]/A$ donde A es algún ideal de $\mathbb{F}_q[x]$. Y recordemos que $\mathbb{F}_q[x]$ es un anillo de ideales principales lo que significa que podemos representar A como $A = (f(x))$.
Si usamos las operaciones de los polinomios para encontrar el primer cambio cíclico, esto significaria que queremos mover el coeficiente a_0 a el coeficiente de x, a_1 a el coeficiente de x^2 y así sucesivamente, es decir claramente debemos multiplicar por x:

$$\begin{aligned} a(x)x &= (a_0 + a_1x + a_2x^2 + ... + a_{n-2}x^{n-2} + a_{n-1}x^{n-1})x \\ &= a_0x + a_1x^2 + a_2x^3 + ... + a_{n-2}x^{n-1} + a_{n-1}x^n \\ &= a_{n-1}x^n + a_0x + a_1x^2 + a_2x^3 + ... + a_{n-2}x^{n-1} \end{aligned}$$

Este resultado corresponde a $a_{n-1}a_0, ..., a_{n-2}$ solo sí $x^n = 1$. Se sigue que $x^n - 1 = 0$, por lo que todos los polinomios se ven módulo $x^n - 1$, Por lo tanto $f(x) = x^n - 1$ y $R_n = \mathbb{F}_q[x]/(x^n - 1)$

Proposición 10.40. *Un conjunto de elementos S en R_n corresponde a un código cíclico C si y sólo si S es un ideal de R_n.*

Demostración. Supongamos que S es un conjunto de elementos de R_n que corresponde a un código ciclico. Entonces si $a_1(x)$ y $a_2(x)$ están en S, por la definición de un código $a_1(x) \pm a_2(x)$ esta en S. Recordemos que el cambio ciclico corresponde a la multiplicación por x entonces si $a(x)$ esta en S, entonces $a(x)x$ esta en S. Consideremos $a(x)$ en S y

$$b(x) = \sum_{i=0}^{n-1} b_i x^i$$

donde $b(x)$ es algún polinomio en R_n luego

$$a(x)b(x) = \sum_{i=0}^{n-1} b_i a(x) x^i$$

Pero cada elemento de la suma es un elemento de S, por lo tanto $a(x)b(x)$ esta en S. Por lo tanto S es un ideal. Si S es un ideal en R_n entonces los polinomios en S corresponde a los vectores en un código ciclico. □

R_n es un anillo de ideales principales, por lo que un código ciclico C le correspondera el ideal $(g(x))$ para algún $g(x) \in R_n$.
En los próximos teoremas daremos información acerca de $g(x)$.

Proposición 10.41. *Si C es un ideal de $R_n = \mathbb{F}_q[x]/(x^n - 1)$, el cual por el Teorema 10.40 lo convierte en un código ciclico. Sea $g(x)$ un polinomio mónico con el grado más pequeño en C. Entonces $g(x)$ es el único polinomio mónico con el grado más pequeño y $C = (g(x))$.*

Demostración. Veamos que R_n es un anillo de ideales principales y que el generador mónico de grado más pequeño de un ideal es único aunque un ideal puede tener otros generadores.
Primero vamos a mostrar que R_n es un anillo de ideales principales. Sea $g(x)$ el polinomio mónico de grado más pequeño en $C \neq (0)$, y sea $a(x)$ otro polinomio en C. Por el algoritmo de la división en $\mathbb{F}_q[x]$, $a(x) = g(x)b(x) + r(x)$ donde $\mathbf{gr}(r(x)) < \mathbf{gr}(g(x))$. Por la definición de un ideal $r(x)$ esta en C pero esto contradice el hecho de que tomamos a $g(x)$ como el polinomio de grado más pequeño, así $r(x) = 0$ y entonces $a(x) = g(x)b(x)$. Por lo tanto R_n es anillo de ideales principales.
Si $g(x)$ y $h(x)$ son polinomios mónicos de el mismo grado y ambos están en C, entonces $g(x) - h(x) \in C$ es de menor grado. Esto no puede pasar si $g(x)$ tiene el menor grado. Por lo tanto $g(x)$ es el único polinomio mónico de grado más pequeño en C y $C = (g(x))$ □

El siguiente teorema explica como encontrar este generador de un código cíclico.

Proposición 10.42. *Si C es un ideal en R_n, entonces su único generador mónico $g(x)$ divide $x^n - 1$. Inversamente, si $g(x) \in C$ y divide $x^n - 1$. Entonces $g(x)$ tiene el grado más pequeño en $(g(x))$.*

Demostración. Supongamos primero que $g(x)$ es el polinomio mónico de grado más pequeño en C. Por el algoritmo de la división en $\mathbb{F}_q[x]$, $x^n - 1 = a(x)g(x) + r(x)$ donde el $\mathbf{gr}(r(x)) < \mathbf{gr}(g(x))$. Ahora $r(x) = -a(x)g(x)$ módulo $(x^n - 1)$, y entonces $r(x)$ está en $(g(x))$. Está es una contradicción a no ser que $r(x) = 0$. Por lo tanto $g(x)$ divide $x^n - 1$.
Inversamente, supongamos que $g(x)$ divide $x^n - 1$ y que $b(x)$ esta en $(g(x))$ pero tiene menor grado que $g(x)$. Entonces $b(x) = c(x)g(x) + (x^n - 1)d(x)$ en $\mathbb{F}_q[x]$ porque $b(x)$ está en C. Sin embargo ya que $g(x)$ divide $x^n - 1$, $g(x)$ divide $b(x)$, el cual es una contradicción. □

Juntos, estos teoremas determinan todos los códigos ciclicos.
Del Teorema 10.40 sabemos que todos los códigos cíclicos son ideales de R_n. El teorema 10.41 dice que podemos corresponder un ideal (y por lo tanto un código) con un polinomio en R_n, finalmente el Teorema 10.42 dice exactamente cuales polinomios mónicos generan ideales, que son aquellos que dividen $x^n - 1$.

Observación 10.43. *El polinomio mónico de grado más pequeño $g(x)$ en C es llamado el* ***polinomio generador*** *de C*

Proposición 10.44. *Si C corresponde a $(g(x))$ donde $g(x) = g_0 + g_1x + ... + g_{n-k}x^{n-k} \in R_n$ y el* ***gr***$(g(x))$ *es $n-k$, entonces la dimensión de C es k y una matriz generadora es:*

$$\begin{pmatrix} g_0 & g_1 & g_2 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k-1} & g_{n-k} & \cdots & 0 \\ & \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & & & & g_{n-k} \end{pmatrix}$$

Demostración. Esto es equivalente a mostrar que los vectores

$$g(x), g(x)x, g(x)x^2, ..., g(x)x^{k-1}$$

en R_n son linealmente independientes y generan C. Supongamos que ellos no son linealmente independientes. Entonces hay una combinación lineal de esos vectores con algunos coeficientes distintos de cero el cual es igual a cero:

$$a_0(g(x)) + a_1(g(x)x) + ... + a_{k-1}(g(x)x^{k-1}) = (a_0 + a_1x + ... + a_{k-1}x^{k-1})g(x) = 0$$

Pero el $\mathbf{gr}(g(x)) = n - k$, entonces el producto es un polinomio de grado $k - 1 + n - k = n - 1$, entonces el polinomio no puede ser 0 módulo $x^n - 1$ a menos que todos los a_i sean 0. Para ver que los vectores generan, recordemos que el código es un ideal generado por $g(x)$, entonces todo polinomio puede ser escrito $c(x)g(x) = c_0g(x) + c_1xg(x) + ... + c_{k-1}x^{k-1}g(x)$. □

Lema 10.45. *Sea R un anillo con ideal I y un elemento idempotente $a \in I$. El elemento a es un generador de I si y sólo si actua como la unidad en I.*

Demostración. Supongamos que a es un generador de I. Sea $c = ba$ para algún b en R. Ahora $ca = (ba)a = b(a^2) = ba = c$. Entonces para elementos de I, a actua como la unidad.

Supongamos a actua como la unidad de I y $c \in I$. Para mostrar que $I = (a)$ debemos mostrar que $c = ba$ para algún b. Pero esto es tan simple como $c = ca$ porque a es la unidad en I. Por lo tanto a genera I. □

Finalmente, hablemos del código dual de un código ciclico. Notemos que si $g(x)$ es el polinomio generador de algún código C entonces $x^n - 1 = g(x)h(x)$ para algún $h(x)$ porque $g(x)$ divide $x^n - 1$. Ahora $h(x)$ es llamado el polinomio checador de C, aunque no necesariamente genera el código dual.

Definición 10.46. *Consideremos un número s tal que $0 \leq s \leq p^m - 1$. Sea r el número más pequeño tal que $p^{r+1}s \equiv s (mod\ p^m - 1)$. La clase ciclotómica de s es $\{s, ps, p^2s, ..., p^r s\}$ donde los elementos son vistos módulos $p^m - 1$.*

Proposición 10.47. *Las clases ciclotómicas particionan el conjunto*

$$\{0, 1, 2, ..., p^m - 1\}.$$

Demostración. Definamos la relación $\sim$ sobre el conjunto tal que $x \sim y$ si $x \equiv p^k y (\text{mod } p^m - 1)$ para algún entero k donde $0 \leq k \leq m - 1$. Veamos que esto es una relación de equivalencia, el cual implica que particiona el conjunto en clases de equivalencias. Es reflexiva porque $x \equiv p^0 x$. Es simétrica porque si $x \sim y$, entonces $x \equiv p^k y$, lo que implica $p^{m-k}x \equiv p^{m-k}p^k y \equiv p^m y \equiv y$, lo que implica que $y \equiv p^{m-k}x$, es decir $y \sim x$. Finalmente es transitiva. Si $x \sim y$ y $y \sim z$, entonces $x \equiv p^k y$ y $y \equiv p^l z$. Entonces $x \equiv p^k(p^l z) \equiv p^{k+l} z$. Y observemos que $k + l$, puede ser más grande que $m - 1$ pero no es tan grande como $2m - 2$. Si $k + l > m - 1$, simplemente factorizamos p^m, y notemos que $p^m \equiv 1 (\text{mod } p^m - 1)$. Entonces $x \equiv p^{k+l} z \equiv p^m p^{k+l} p^{-m} z \equiv p^{k+l-m} z$ lo que significa que $x \sim z$ lo que prueba la transitividad. □

Capítulo 11

Anillos Conmutativos Finitos

En está sección estudiaremos algunos resultados interesantes de los anillos finitos y los Anillos de Galois entre los cuales está que un anillo finito se puede expresar como la suma directa de anillos locales y el Lema de Hensel que tiene bastantes aplicaciones en la Teoría de códigos sobre anillos de Galois.

11.1. Estructura de Anillos finitos conmutativos

Sean $I_1, I_2, I_3, ..., I_n$ ideales propios de un anillo R; I_j y I_k, con $1 \leq j \neq k \leq n$ se dice que son **ideales primos relativos** o **coprimos** si $I_j + I_k = R$ donde

$$I_j + I_k := \{a + b \mid a \in I_j \text{ y } b \in I_k\}$$

Definición 11.1. *Sean $\{R_i\}_{i\in A}$ anillos conmutativos, consideremos el **producto infinito de anillos** como sigue:*

$$\prod_{i\in A} R_i = \{f : A \to \bigcup_{i\in A} R_i \mid f(i) \in R_i \text{ para cada } i \in A\}$$

Proposición 11.2. *$(\prod_{i\in A} R_i, +, \hat{0}, \cdot, \hat{1})$ es un anillo conmutativo.*

Demostración. Sean $f, g \in \prod_{i\in A} R_i$, definimos $(f + g) : A \to \bigcup_{i\in A} R_i$ como $(f + g)(i) = f(i) + g(i)$.

1. La función cero, $\hat{0} \in \prod_{i\in A} R_i$ y es tal que $\hat{0}(i) = O_{R_i}$.

2. $(f + \hat{0})(i) = f(i) + \hat{0}(i) = f(i) + 0_{R_i} = f(i)$.

3. Para $f \in \prod_{i\in A} R_i$, definimos $(-f) : A \to \bigcup_{i\in A} R_i$ como $(-f)(i) = -f(i)$ y es tal que $f + (-f) = \hat{0}$.

4. Sean $f, g, h \in \prod_{i\in A} R_i$

$$\begin{aligned}[(f + g) + h](i) &= (f + g)(i) + h(i)\\ &= f(i) + g(i) + h(i)\\ &= f(i) + (g(i) + h(i))\\ &= f(i) + [g + h](i)\end{aligned}$$

Por lo tanto $(\prod_{i\in A} R_i, \hat{0}, +)$ es un grupo abelino.
Para $f, g \in \prod_{i\in A} R_i$ definimos $(f \cdot g)(i) = f(i) \cdot g(i)$.

1. $\hat{1} : A \to \bigcup_{i\in A} R_i$ y es tal que $\hat{1}(i) = 1_{R_i}$

Por lo tanto $(\prod_{i\in A} R_i, \cdot, \hat{1})$ es un monoide abeliano.

$$\begin{aligned}(f \cdot (g + h))(i) &= f(i)(g + h)(i)\\ &= f(i)(g(i) + h(i))\\ &= f(i)g(i) + f(i)h(i).\end{aligned}$$

Así que $(\prod_{i\in A} R_i, +, \hat{0}, \cdot, \hat{1})$ es un anillo conmutativo. □

Definición 11.3. *Si $f \in \prod_{i\in A} R_i$, el $sop(f) = \{i \in A \mid f(i) \neq 0\}$. Definimos la* ***suma directa*** *como*

$$\bigoplus_{i\in A} R_i = \{f \in \prod_{i\in A} R_i \mid |sop(f)| < \infty\}.$$

Observación 11.4. *Para A finito, es decir $|A| = n$*

$$\prod_{i\in A} R_i = \prod_{i=1}^{n} R_i.$$

Observación 11.5. $\prod_{i=1}^{n} R_i = \{f : \{1, ...n\} \to \bigcup_{i=1}^{n} R_i \mid f(i) \in R_i\}$.

1. $\hat{1} \in \prod_{i=1}^{n} R_i$, $\hat{1}(i) = 1$, *para todo* $i \in \{1, ...n\}$.

2. *Si* $f \in \prod_{i=1}^{n} R_i$, $f(i) \in R_i$, *para cada* $i \in A$. $(f(i))_{i\in A} = f$.

3. *Si* $f(i) = x_i$, $f = (f(i))_{i\in A} = (x_i)_{i\in A}$.

Definición 11.6. *Sean* $I_1, ..., I_n$ *ideales de* R, *entonces* R/I_j *es un anillo para* $j \in \{1, ..., n\}$ *definimos*

$$\prod_{j=1}^{n} R/I_j$$

como el ***anillo producto****.*

Observación 11.7. *Podemos definir un morfismo de anillos de* R *a* $\prod_{i=1}^{n} R/I_j$ *de la siguiente forma*

$$f : R \to \prod_{i=1}^{n} R/I_j$$
$$x \mapsto f(x) = (x + I_j)_{j\in A}$$

con $A = \{1, ..., n\}$.
Veamos que f *es un morfismo de anillos.*

1. *Sean* $x, y \in R$

$$\begin{aligned} f(x+y) &= [(x+y) + I_j]_{j\in A} \\ &= [(x + I_j) + (y + I_j)]_{j\in A} \\ &= (x + I_j)_{j\in A} + (y + I_j)_{j\in A} \\ &= f(x) + f(y). \end{aligned}$$

2. *Sean* $x, y \in R$

$$\begin{aligned} f(xy) &= [(xy) + I_j]_{j\in A} \\ &= [(x + I_j)(y + I_j)]_{j\in A} \\ &= (x + I_j)_{j\in A}(y + I_j)_{j\in A} \\ &= f(x)f(y). \end{aligned}$$

3. $f(1_R) = (1 + I_j)_j j \in A = \hat{1}_{\prod_{i=1}^{n} R/I_j}$

Por lo tanto f *es un morfismo de anillos.*

Proposición 11.8. *f es inyectivo si y sólo si $\bigcap_{j=1}^{n} I_j = \{0\}$.*

Demostración. Recordando la Proposición 3.17. Basta demostrar que $kerf = \bigcap_{j=1}^{n} I_j$.

1. Claramente $\bigcap_{j=1}^{n} I_j \subseteq kerf$.

2. Sea $x \in kerf$, entones $f(x) = \hat{0}_{\prod_{i=1}^{n} R/I_j} = (I_j)_{j\in A}$, por otro lado $f(x) = (x + I_j)_{j\in A}$. Así que $(x + I_j)_{j\in A} = (I_j)_{j\in A}$, es decir, $x + I_j = I_j$ para toda $j \in A$, entonces $x \in I_j$ para toda $j \in A$, por lo tanto $x \in \bigcap_{j=1}^{n} I_j$.

□

Proposición 11.9. *f es sobreyectiva si y sólo si $I_l + I_k = R$, con $l \neq k$.*

Demostración. $\Rightarrow$] Supongamos que f es sobreyectiva. Sean $I_l, I_k \leq R$ con $l \neq k$ y $l, k \in A$. Sea $\delta_l \in \prod_{j=1}^{n} R/I_j$ tal que $\delta_l = (x_i)_{i\in A}$ donde $x_i = 0 + I_i$ si $i \neq l$ y $x_i = 1 + I_i$ si $i = l$. Como $\delta_l \in \prod_{j=1}^{n} R/I_j$ y f es sobreyectiva, entonces existe $x \in R$ tal que $f(x) = \delta_l$, pero $f(x) = (x + I_i)_{i\in A} = (x_i)_{i\in A}$. Entonces $x + I_l = 1 + I_l$ y $x + I_k = 0 + I_k$ con $k \neq l$. De modo que $1 - x \in I_l$ y $x \in I_k$, para todo $k \neq l$. Así $1 = (1 - x) + x \in I_l + I_k$, con $l \neq k$. Por lo tanto $R = I_k + I_l$, es decir, I_l, I, k son coprimos.
$\Leftarrow$] Supongamos que $R = I_k + I_l$ para todo $k \neq l$. Sea $\delta_l = (x_i)_{i\in A}$ donde $x_i = 0 + I_i$ si $i \neq l$ y $x_i = 1 + I_i$ si $i = l$. Veamos que existe $x \in R$ tal que $f(x) = \delta_l$. Entonces $u_k + v_k = 1$, para todo $k \neq l$, $u_k \in I_l$, $v_k \in I_k$, entonces

$$\prod_{k\neq l} v_k = \prod_{k\neq l}(1 - u_k), \tag{11.1}$$

luego $1 - u_k \equiv 1 \bmod I_l$ entonces $\prod(1 - u_k) \equiv 1 \bmod I_l$. Por lo tanto

$$\prod_{k\neq l} v_k \equiv 1 \bmod I_l,$$

luego $\prod_{k\neq l} v_k \in R$, implica $f(\prod_{k\neq l} v_k) = (\prod_{k\neq l} v_k + I_j)_{j\in A}$.

Si $j = l$, $\prod_{k\neq l} v_k + I_l = 1 + I_l$.

Si $j \neq l$, $v_j \in \{v_k\}_{k\neq l}$, con $v_j \in I_j$.

Por lo tanto

$$\prod_{k\neq l} v_k + I_j = I_j = 0 + I_j.$$

Es decir $f(\prod_{k\neq l} v_k) = \delta_l$.

□

Proposición 11.10. *Sea R un anillo finito.*

1. *Si I_j y I_k, $1 \leq j \neq k \leq n$ son ideales primos relativos de R entonces:*

$$\bigcap_{j=1}^{n} I_j = \prod_{j=1}^{n} I_j.$$

2. *Si I_j y I_k son primos relativos, entonces así lo son I_j^m y I_k^m para todo $m \in \mathbb{N}$*
   (Notemos que, si I es un ideal de R, I^m es la m-ésima potencia, es decir el ideal generado por los elementos $x_1 \cdots x_m$ donde $x_k \in I$, $1 \leq k \leq m$).

Demostración. 1. La demostración se hará por inducción sobre n.
Para $n = 2$. Sean $I_1, I_2 \leq R$ tales que $I_1 + I_2 = R$. Es claro que $I_1 I_2 \subseteq I_1 \cap I_2$. Sea $y \in I_1 \cap I_2$, como $I_1 + I_2 = R$, entonces $1 \in R$, entonces $1 = a_1 + a_2$, con $a_1 \in I_1$ y $a_2 \in I_2$, multiplicando por y en ambos lados $y = a_1 y + a_2 y$ con $a_i y \in I_1 I_2$ para $i = 1, 2$. Por lo tanto $y \in I_1 I_2$.
Fijemos ahora a un ideal de R. Sea $I_n \leq R$, y para todo $j \in \{1, ..., n\}$, $I_j + I_n = R$ con $j \neq n$. Así para toda $j \neq n$, $a_j + b_j = 1$ con $a_j \in I_j$ y $b_j \in I_n$, por lo tanto

$$\prod_{j\neq n} a_j = \prod_{j\neq n} (1 - b_j) \equiv 1 \bmod I_n$$

Ya que $1 - b_j = a_j$ y $-b_j = (1 - b_j) - 1 \in I_n$, entonces $1 - b_j \equiv 1 \bmod I_n$ para todo $j \neq n$, así

$$\prod_{j\neq n} (1 - b_j) \equiv \prod (1) = 1.$$

Sea

$$\prod_{j=1}^{n} I_j = \prod_{j=1}^{n-1} I_j I_n = \bigcap_{j=1}^{n-1} I_j I_n = K I_n$$

Por demostrar que $K + I_n = R$.
Para todo $j \neq n$, $\prod_{j\neq n} a_j \in I_j$ y $\prod_{j\neq n} a_j \in \bigcap_{j=1}^{n-1} I_j$, así $1 - \prod_{j\neq n} a_j \in I_n$, entonces $1 - \prod_{j\neq n} a_j = b$, para $b \in I_n$, entonces $1 = \prod_{j\neq n} a_j + b$, es decir, $K + I_n = R$. Por lo tanto $KI_n = K \cap I_n = \bigcap_{j=1}^{n-1} I_j I_n = \bigcap_{j=1}^{n} I_j$.

2. Por hipótesis, I_j y I_k son primos relativos, así existen $x_j \in I_j$ y $x_k \in I_k$ tal que $x_j + x_k = 1$. Esto significa que $1 = (1)(1) = (x_j + x_k)(x_j + x_k) = x_j^2 + x_k^2 + 2x_j x_k$; Hay dos posibilidades: Si $x_j x_k = 0$, entonces inmediatamente se sigue que $R = I_j^2 + I_k^2$. De otra manera $2x_j x_k = (2x_j + 2x_k)x_j x_k = 2x_j^2 x_k + 2x_j x_k^2 \in I_j^2 + I_k^2$. Por lo tanto, por lo anterior $1 \in I_j^2 + I_k^2$.

□

Definición 11.11. *Un elemento e de un anillo R es llamado idempotente si $e^2 = e$. Dos elementos idempotentes de R, e y f, se dicen ortogonales si $ef = 0$.*

Ejemplo 11.12. *Veamos $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$*
Los elementos idempotentes de $\mathbb{Z}_6$ son 0,1,3,4, mientras que los elementos ortogonales son 3 y 4.

Proposición 11.13. *Sea R un anillo finito. Las siguientes proposiciones son equivalentes:*

1. R es isomorfo a una suma directa de subanillos R_j , $1 \leq j \leq n$.

2. Existen elementos idempotentes ortogonales e_j tales que

$$1 = \sum_{i=1}^{n} e_j \text{ y } R_j \cong Re_j$$

3. R es una suma directa de ideales propios $I_j \cong R_j$, $1 \leq j \leq n$

Demostración. 1. $\Rightarrow$]2. Por hipótesis existen $e_j \in R_j$, para todo $j \in \{1, ..., n\}$, tal que $1 = \sum_{j=1}^{n} e_j$. Si consideramos e_k como un elemento del anillo R, entonces $e_k = \sum_{j=1}^{n} e_k e_j$ lo cual significa $e_k e_j = \delta_{kj} e_k$, donde δ_{kj} es la delta de Kronecher, lo que significa que las $e_j's$ donde $1 \leq j \leq n$ son elementos idempotentes ortogonales de R. Más aún, R_j es el ideal principal de R generado por e_j
2. $\Rightarrow$ 3.
Por el paso anterior, todo R_j es un ideal de R.
3. $\Rightarrow$ 1. Es claro. □

Proposición 11.14. *Un anillo finito R puede ser expresado como una suma directa de anillos locales.*

Demostración. Sean $P_1, P_2, ..., P_n$ ideales primos de R, es decir $Spec(R) = \{P_1, ..., P_n\}$. Por las Proposiciones 6.8 y 1.17, R/P_i es un campo, luego por la Proposición 6.6 cada P_i con $i = \{1, ..., n\}$ es ideal máximo de R, es decir $spec(R) = specm(R)$. De modo que, el radical de Jacobson coincide con el nilradical de R. Por ser máximos los $P_i's$, $1 \leq i \leq n$, se tiene que todo par de ideales (P_j, P_k), $1 \leq j \neq k \leq n$, es un par de ideales coprimos de R, así $\bigcap_{j=1}^{n} P_j = \prod_{j=1}^{n} P_j$. Ya que $J(R)$ es un ideal nilpotente existe $m_0 \in \mathbb{N}$ tal que $J(R)^{m_0} = \{0\}$. Definimos el morfismo de anillos.

$$\phi_0 : R \to R/P_1^{m_0} \oplus ... \oplus R/P_n^{m_0} \tag{11.2}$$

de forma natural. Lo que probamos en las Proposiciones 11.10, 11.8 y 11.9, nos asegura que ϕ_0 sea un isomorfimo, ya que cualquiera dos ideales $P_j^{m_0}$, $1 \leq j \leq n$, son coprimos y $ker(\phi_0) = \bigcap_{j=1}^{n} P_j^{m_0} = \prod_{j=1}^{n} P_j^{m_0} = J(R)^{m_0} = \{0\}$. Este isomorfimos determina una biyección entre los ideales propios del anillo $R/P_j^{m_0}$, $1 \leq j \leq n$ y los ideales propios de R. Ya que P_j es el único ideal máximo de R tal que $P_j^{m_0} < P_j < R$, entonces por el Teorema de la correspondencia $R/P_j^{m_0}$ es un anillo local con ideal máximo $P_j/P_j^{m_0}$. Asumamos que existen dos descomposiciones distintas de R como una suma directa de anillos locales.

$$R = \bigoplus_{j=1}^{n} R_j = \bigoplus_{k=1}^{m} S_k.$$

Entonces existen elementos idempotentes ortogonales $e_j \in R_j$ y $f_k \in S_k$, $1 \leq j \leq n$, $1 \leq k \leq m$, tal que

$$1 = \sum_{j=1}^{n} e_j = \sum_{k=1}^{m} f_k.$$

Cada sumando propio R_j es isomorfo a un anillo local Re_j, analogamente, cada S_k es isomorfo a un anillo local Rf_k. De ahí, ningúno de los elementos e_j y f_k es una suma de dos o más elementos idempotentes propios, de hecho, en general, un anillo local no contiene elementos idempotentes diferentes de 0 y 1, ya que su radical de Jacobson es el ideal máximo. Por lo tanto $e_j = \sum_{k=1}^{m} e_j f_k$, así existe un entero k_j tal que $e_j = e_j f_{k_j}$ y, analogamente, existe un entero j_k tal que $f_k = f_k e_{j_k}$, $1 \leq j \leq n, 1 \leq k \leq m$. Esto significa que

$$e_j = e_j f_{k_j} = e_j f_{k_j} e_{j_{k_j}},$$

así $j = j_{k_j}$ como los elementos $\{e_j\}_{1 \leq j \leq n}$ son mutuamente ortogonales. □

Proposición 11.15. *Sea R un anillo y $A = \{r \in R : r$ no es unidad $\}$ Entonces son equivalentes las siguientes condiciones.*

I. *Para todo $r_1, r_2 \in A$, $r_1 + r_2 \in A$.*

II. *A es un ideal.*

III. *A es el ideal mayor propio de R.*

IV. *Existe un ideal mayor propio de R.*

V. *Para todo $r \in R$, r es unidad o $(1 - r)$ es unidad.*

Demostración. i) ⇒ ii) Basta con que veamos que, es cerrado bajo el producto con elementos del anillo, es decir. Para todo $a \in A$ y $r \in R$ entonces $ra \in A$. Supongamos $ra \notin A$, eso significa que ra es unidad, entonces existe $u \in R$ tal que $1 = (ar)u = a(ru)$ lo que implica a es unidad, pero esto es una contradicción. Por lo tanto $ra \in A$.
ii) ⇒ iii) Tomemos B ideal de R. Por demostrar que $B \subseteq A$, supongamos que no se cumple $B \subseteq A$. entonces existe $b \in B$ tal que $b \notin A$. como $b \notin A$ implica que b es unidad entonces existe $u \in R$ tal que $bu = 1$ luego como B es ideal y $b \in B$ se sigue que $1 \in B$.
iii)⇒ iv) Es evidente.
iv)⇒ v) Sea C el ideal mayor propio de R. Sea $r \in R$ y supongamos r y $1-r$ no son unidad, entonces $r \in A$ y $1 - r \in A$, luego $r \in C$ y $1 - r \in C$. Así $1 \in C$, lo que es una contradicción.
v) ⇒ i) Sean $r_1, r_2 \in A$ y supongamos $r_1 + r_2 \notin A$, entonces existe $u \in R$ tal que $(r_1 + r_2)u = r_1u + r_2u = 1$ así que $r_1u = 1 - r_2u$, además por la prueba de i) ⇒ ii) $r_1u \in A$, entonces $1 - r_2u \in A$, pero si $1 - r_2u \in A$ implica que $1 - r_2u$ no es unidad, entonces r_2u es unidad, pero $r_2u \in A$ es decir r_2u no es unidad. Lo que es una contradicción. □

Definición 11.16. *Un anillo R es llamado de Galois si es finito, local y su ideal máximo esta dado por (p) donde p es primo.*

Proposición 11.17. *El anillo $\mathbb{Z}_{p^n}$ es un anillo de Galois.*

Demostración. Mostraremos que (p) es el único ideal máximo. Consideremos el morfismo

$$\begin{aligned} \phi : \mathbb{Z}_{p^n} &\longrightarrow \mathbb{Z}_p \\ a \bmod p^n &\longmapsto a \bmod p \end{aligned}$$

el cual se puede ver que es sobreyectiva. Observemos que $ker(\phi) = (p)$. Usando el primer teorema para isomorfismos para anillos tenemos que

$$\mathbb{Z}_{p^n}/(p) \cong \mathbb{Z}_p.$$

Por la Proposición 6.6, vemos que (p) es máximo. Para mostrar la unicidad, supongamos M fuera otro ideal máximo distinto de (p). Notemos que $p \notin M$. Supongamos que $p \in M$. Entonces $(p) \subseteq M$ con $(p) \neq M$ significa que M contiene propiamente a (p). Ya que (p) es máximo, esto implica $M = \mathbb{Z}_{p^n}$. Un ideal máximo es un ideal propio por definición, pero esto es una contradicción que muestra que $p \notin M$. Nuevamente por la Proposición 6.6, $\mathbb{Z}_{p^n}/M$ debe ser un campo. Ya que M es un ideal propio, $1 \notin M$ y entonces $1 + M \neq 0$. Ya que todos los campos tienen característica un primo, en este caso p, esto implica $p1 + M = M$ de modo que $p \in M$. Esta contradicción prueba la unicidad de (p). Por lo que $\mathbb{Z}_{p^n}$ es un anillo de Galois.

□

11.2. Propiedades importantes de anillos de Galois

Lema 11.18. *Sea R un anillo finito, local con un único ideal máximo M. Si I es un ideal propio de R, entonces $I \subseteq M$.*

Demostración. Supongamos que I no es un subconjunto de M. Entonces hay algún $a \in I$ tal que $a \notin M$. Por lo tanto $(a) \not\subseteq M$. Si (a) es máximo, entonces contradice la unicidad de M . Entonces debe haber algún ideal A_1, tal que $(a) \subseteq A_1$, si A_1 es máximo, nuevamente contradice la unicidad de M. Continuando inductivamente debe haber siempre un ideal más grande. Pero R es finito, entonces esta cadena de ideales debe ser finita, es decir hay algún ideal A_r final, el cual otra vez contradice la unicidad de M. Por lo tanto $I \subseteq M$ □

Proposición 11.19. *Sea R un anillo de Galois cuyos divisores de cero junto con 0 forma un ideal principal $(p1)$ para un número primo p. Entonces $(p1)$*

es el único ideal máximo de R, $R/(p1)$ es un campo de característica p. La característica de R es una potencia de p.

Demostración. En un anillo finito todo elemento distinto de cero el cual no es un divisor de cero es una unidad. Por lo tanto $(p1)$ es el único ideal máximo de R. Denotamos el morfismo sobreyectivo natural $f : R \to R/(p1)$ tal que $f(r) = r + (p1)$. Entonces $p(1 + (p1)) = p + (p1) = 0 + (p1)$. De ahí $R/(p1)$ es un campo finito de característica p. Sea k la característica de R, es decir, $k1 = 0$, luego

$$\begin{aligned} 0 + (p1) &= f(k1) \\ &= k + (p1) \\ &= (k + (p1))(1 + (p1)). \end{aligned}$$

De ahí $p \mid k$. Asumamos que $k = p^s l$, donde s, l son enteros positivos y su $m.c.d(p, l) = 1$. Si $l > 1$, entonces $a = p^s l$ y $b = l1$ elementos distintos de cero en R y $ab = 0$. Así $l1 \in (p1)$ y $l(1 + (p1)) = l + (p1) = 0 + (p1)$. Pero $R/(p1)$ es de característica p, así $p \mid l$, lo que es una contradicción, ya que $m.c.d(p, l) = 1$. Por lo tanto $l = 1$ y $k = p^s$.

□

Proposición 11.20. *Sea R un anillo finito, local con un único ideal máximo M. Un elemento $x \in R$ es una unidad si y sólo si $x \notin M$.*

Demostración. La demostración se hará por contrarecíproca. Supongamos que $x \in M$, deseamos mostrar que x no es unidad. Observemos que $(x) \subseteq M$. Por lo tanto para todo $r \in R$ tenemos que $rx \in M$. Ahora $1 \notin M$ porque de otra forma M no sería un ideal propio. De ahí no podemos tener un elemento $y \in M$ tal que $xy = 1$. Recíprocamente, supongamos que $x \notin M$. Entonces el ideal $(x) \not\subseteq M$. Por el Lema 11.18, sin embargo, M contiene todos los ideales propios de R. Por lo tanto, debemos tener $(x) = R$, lo que implica que $1 \in (x)$, así hay un $y \in R$ tal que $xy = 1$. □

La proposición 11.20 nos dice como deben ser las unidades de un anillo de Galois.

Lema 11.21. *El conjunto de los elementos nilpotentes forman un ideal.*

Demostración. Notemos que N es distinto del vacío ya que $0 \in N$. Sea $N = \{a \in R \mid a^k = 0, \text{ para algún } k \in \mathbb{N}\}$. Entonces si $r \in R$ y $a \in N$, tenemos $(ra)^k = r^k a^k = r^k 0 = 0$. Ahora supongamos que $a, b \in N$, entonces existen enteros n y m tal que $a^n = b^m = 0$. Consideremos $a^j b^{n+m-j}$, si $j \geq n$ entonces $a^j = 0$ así $a^j b^{n+m-j} = 0$, mientras si $0 \leq j < n$ entonces $b^{n+m-j} = 0$ así $a^j b^{n+m-j} = 0$. Por lo tanto por el Teorema del Binomio se tiene que $(a+b)^{n+m} = 0$, así $a + b \in N$. □

Como se vio en el Lema 11.18, $N \subseteq M$. Veamos que $M \subseteq N$, mostrando que $N = M$.

Proposición 11.22. *El conjunto de elementos nilpotentes en R es exactamente el ideal maximal M.*

Demostración. Sea N es el ideal de elementos nilpotentes. Ya sabemos que $N \subseteq M$. Por demostrar que $M \subseteq N$. Ya que $M = (p)$, esto es equivalente a mostrar que si un elemento de R es divisible por p , entonces este también es nilpotente. Recordemos que la característica de R es p^k para algún $k \in \mathbb{N}$, así $p^k = 0$, mostrando que p es nilpotente. Entonces si a es divisible por p, tenemos $a = bp$. Esto implica que a es nilpotente, ya que

$$a^k = (bp)^k = b^k p^k = b^k 0 = 0$$

□

Otro hecho útil es

Proposición 11.23. *Para todo anillo R, si u es una unidad y a es nilpotente, entonces $u + a$ es una unidad.*

Demostración. El inverso de $u + a$ esta dado por $(u^{k-1} - u^{k-2}a + ... + (-1)^{k-1}a^{k-1})(u^{-1})^k$ donde k es tal que $a^k = 0$.

$$(u+a)(u^{k-1} - u^{k-2}a + ... + (-1)^{k-1}a^{k-1})(u^{-1})^k$$

$$= (u^k + u^{k-1}a - u^{k-1}a + ... + (-1)^{k-2}ua^{k-1} + (-1)^{k-1}ua^{k-1} + (-1)^{k-1}a^k)(u^{-1})^k$$

$$= (u^k + (-1)^{k-1}a^k)(u^{-1})^k$$

$$= (u^k)(u^{-1})^k = 1$$

□

Con frecuencía consideraremos anillos de polinomios sobre anillos de Galois, por lo que debemos estudiar algunas propiedades.
Con la Proposición 11.23 se puede demostrar lo siguiente :

Proposición 11.24. *Sea $f(x) = a_+a_1x + ... + a_kx^k$ un polinomio en $R[x]$. Entonces $f(x)$ es una unidad en $R[x]$ si y sólo si $a_0 \notin M$ y $a_1, ..., a_k \in M$.*

Demostración. Supongamos que $f(x)$ es unidad. Y sea $\psi : R[x] \to R[x]/M$ un morfismo que toma los coeficientes de polinomios a sus correspondientes clases en R/M. Ya que R/M es un campo, las unidades de $R[x]/M$ son los polinomios constantes $a + M$. Ahora si $f(x)$ es una unidad en $R[x]$, entonces $\psi(f(x))$ es también una unidad. Esto es porque si $f(x)g(x) = 1$, entonces $1 + M = \psi(1) = \psi(f(x)g(x)) = \psi(f(x))\psi(g(x))$. Usando este hecho $\psi(f(x))$ sigue siendo una unidad en $R[x]$, significa $\psi(f(x)) = \psi(a_0) + \psi(a_1)x + ... + \psi(a_k)x^k = a_0 + ... + a_kx^k + M$, es constante y distinto de cero. Por lo tanto, $a_0 \notin M$ y el resto de los coeficientes están en M.
Ya que $a_0 \notin M$ esto significa que a_0 es unidad en R , así como en $R[x]$. El elemento $a_1x + ... + a_kx^k$ es nilpotente en $R[x]$. Ya que $R[x]$ es un anillo unitario, la Proposición 11.23 implica que $f(x) = a_0 + a_1x + ... + a_kx^k$ es una unidad también. □

Proposición 11.25. *$f(x) \in R[x]$ es un elemento nilpotente en $R[x]$ si y sólo si $a_0, ..., a_n$ son nilpotentes.*

Demostración. $\Rightarrow$] Como $f(x)$ es nilpotente, $1 + f(x)$ es una unidad en $R[x]$. Por la Proposición 11.24, $a_1, ..., a_n$ son nilpotentes en R mientras $1 + a_0 \in U(R)$. Por lo tanto para n lo suficientemente grande, $f^n = 0$, implica que $a_0^n = 0$, así a_0 es nilpotente también.
$\Leftarrow$] Si $n_j \in \mathbb{N}$ es tal que $a_j^{n_j} = 0$, $0 \leq j \leq n$ y $n_j \geq 2$, definiendo

$$\bar{n} := (\textstyle\sum_{j=0}^{n} n_j) - n$$

tenemos $f(x)^{\bar{n}} = 0$. De hecho, $f(x)^{\bar{n}}$ es una combinación lineal, con coeficientes enteros, de productos de la forma

$$a_0^{r_0} a_1^{r_1} ... a_t^{r_t} ... a_n^{r_n} x^{k_t}$$

tal que $\sum_{j=0}^{n} r_j = \bar{n}$, para cada $0 \leq k_t < n\bar{n}$.
Ya que no podemos simultaneamente tener $r_j < n_j$, para cada j, cada uno de estos productos es cero. □

Proposición 11.26. *Sea R un anillo. $f(x) \in R[x]$ es un divisor de cero si y sólo si existe un elemento $0 \neq a \in R$ tal que $af(x) = 0$.*

Demostración. $\Rightarrow$] Escojamos un polinomio particular $g \in \{h \in R[x] \mid h \neq 0, hf = 0\} \neq \emptyset$ con grado mínimo. Digamos $g(x) = b_0 + b_1x + ... + b_mx^m$ con $b_m \neq 0$. Afirmamos que $b_mf = 0$. De lo contrario existe un entero $r \in \{0, 1, ..., n\}$ tal que $b_ma_r \neq 0$ así que $b_ma_i = 0$, para todo $i \in \{r + 1, ..., n\}$. Entonces,para todo $i \in \{r + 1, ..., n\}$, tenemos $a_igf = 0$ y el grado de

$$\begin{aligned} a_ig &= a_i(b_0 + b_1x + ... + b_mx^m) \\ &= a_ib_0 + a_ib_1x + ... + a_ib_{m-1}x^{m-1} \end{aligned}$$

es menor que el grado de g. Por nuestra elección de g, sabemos que $a_ig = 0$ para todo $i \in \{r + 1, ..., n\}$. Pero entonces tenemos

$$\begin{aligned} 0 = fg &= (a_0 + ... + a_rx^r + a_{r+1}x^{r+1} + ... + a_nx^n)g \\ &= (a_0 + ... + a_rx^r)g \\ &= (a_0 + ... + a_rx^r)(b_0 + ... + b_mx^m), \end{aligned}$$

lo cual fuerza $a_rb_m = 0$, lo cual es una contradicción.

$\Leftarrow$] Se cumple por definición. $\square$

Capítulo 12

Polinomios regulares en el anillo $R[x]$

En esta sección seguiremos con nuestro estudio del anillo de polinomios pero esta vez vamos a considerar a R como un anillo finito, local y conmutativo con ideal máximo M único y campo residual $K = R/M$.
El $\pi : R \to K$ morfismo sobreyectivo natural se extiende a un morfismo de anillos de polinomios.

$$\mu : R[x] \to K[x]$$

Definición 12.1. *Sea A un anillo conmutativo, un ideal I de A es llamado* ***primario*** *si $I \neq A$ además cuando $xy \in I$ y $x \notin I$ entonces $y^n \in I$, para algún $n \in \mathbb{N}$.*

Definición 12.2. *Sea f y g elementos de $R[x]$*

1. *f es regular si no es un divisor de cero.*

2. *f es primario si (f) es un ideal primario.*

3. *f y g son primos relativos si $R[x] = (f) + (g)$.*

Proposición 12.3. *Sea $f(x) = a_0 + a_1x + ... + a_nx^n$ un elemento de $R[x]$, las siguientes condiciones son equivalentes:*

1. *f es una unidad.*

2. *$\mu(f)$ es una unidad en $K[x]$.*

3. a_0 es una unidad en R y $a_1, ..., a_n$ son nilpotentes.

Demostración. 1. $\Rightarrow$ 2. Si f es una unidad, entonces existe un polinomio g tal que $fg = 1$. Consecuentemente, $1 = \mu(1) = \mu(fg) = \mu(f)\mu(g)$, así $\mu(f)$ es una unidad.
2. $\Rightarrow$ 3. Las únicas unidades en $K[x]$ son los polinomios constantes $\mu(f) = c$, así por definición de μ, los coeficientes a_1, $1 \leq i \leq n$, deben pertenecer a M, es decir son nilpotentes, a_0 es de la forma $a_0 = c + h$, donde h es un elemento nilpotente y c es una unidad; por lo tanto a_0 es invertible.
3. $\Rightarrow$ 1. Es una consecuencia de la Proposición 11.24. □

Proposición 12.4. *Sea $f(x) = a_0 + a_1x + ... + a_nx^n$ un polinomio en $R[x]$ lo siguiente es equivalente:*

1. f es nilpotente.

2. $\mu(f) = 0$.

3. $a_0, ..., a_n$ son nilpotentes en R.

4. f es un divisor de cero.

5. Exite un elemento $a \in R \setminus \{0\}$ tal que $af(x) = 0$.

Demostración. Las implicaciones 2. $\Leftrightarrow$ 3. y 3. $\Leftrightarrow$ 4. se siguen del hecho que R es local y finito, así es suficiente mostrar que 3. es equivalente a 1. y 5.
Por la Proposición 11.25, $f(x)$ es nilpotente si y sólo si sus coeficientes son nilpotentes. La implicación 3. $\Rightarrow$ 5. se sigue de la proposición 11.26 ya que, si $f(x)$ es nilpotente, entonces este es un divisor de cero. Veamos 5. $\Rightarrow$ 3. Supongamos que existe $a \in R \setminus \{0\}$, que por hipótesis cumple $aa_i = 0$ para todo $0 \leq i \leq n$, de manera que todos los a_i son divisores de cero en R; Por lo tanto por la estructura de R, ellos son nilpotentes. □

Proposición 12.5. *Sea $f(x) = \sum_{i=0}^{n} a_ix^i$ un polinomio en $R[x]$. Las siguientes condiciones son equivalentes:*

1. f es regular.

2. El ideal generado por $a_0, a_1, ..., a_n$ coincide con R.

3. a_i es una unidad en R para algún i, $0 \leq i \leq n$.

4. $\mu(f) \neq 0$.

Demostración. 1. $\Rightarrow$ 2. Se sigue de 12.4 (3); de hecho, existe un subíndice $i \in \{1, ..., n\}$ tal que a_i es una unidad en R.
2. $\Rightarrow$ 3. Obvio.
3. $\Rightarrow$ 4. Obvio.
4. $\Rightarrow$ 5. Si $\mu(f) \neq 0$, entonces f no es un divisor de cero en $R[x]$ (ver 12.4 (4)). □

Si A es un ideal de un anillo R, escribimos $A[x]$ para denotar el subanillo de $R[x]$ definido por

$$A[x] := \{a_0 + a_1x + ... + a_nx^n \mid n \geq 0, a_i \in A, 0 \leq i \leq n\}$$

Proposición 12.6. *Sea R finito, local y M su ideal máximo. Entonces*

1. $M[x] = \bigcap_{P \in R[x]} P$, *donde P es un ideal primo en $R[x]$.*

2. $M[x] = \{f(x) \in R[x] \mid g(x)f(x) + 1$ *tiene un inverso, para todo* $g(x) \in R[x]\} = J(R[x])$.

Demostración. 1. Por la Proposición 12.4,

$$M[x] = \{f(x) \in R[x] \mid f(x) \text{ nilpotente}\} = Nil(R[x]).$$

2. Sea $f(x) \in M[x]$; ya que $M[x]$ es un ideal en $R[x]$, $g(x)f(x)$ es nilpotente, para todo $g(x)$ en $R[x]$. De ahí, $M[x] \subseteq J(R[x])$. Por otro lado, si $f(x) \in J(R[x])$, donde $f(x) = \sum_{i=0}^{n} a_ix^i$ con $a_i \in R$, entonces $xf(x) + 1$ tiene un inverso □

Proposición 12.7. *Sea f un elemento de $R[x]$, donde R es un anillo local, finito y sea $\mu(f) = \bar{g_1}...\bar{g_n}$, donde $\bar{g_1}...\bar{g_n} \in K[x]$ son polinomios primos relativos disjuntos en el dominio euclidiano $K[x]$. Entonces existen polinomios $g_1, ..., g_n \in R[x]$ tal que*

1. $g_1, ..., g_n$ *son primos relativos disjuntos en $R[x]$.*

2. $\mu(g_i) = \bar{g_i}$, $1 \leq i \leq n$.

3. $f = g_1, ..., g_n$.

Demostración. Por inducción sobre n. Para $n = 2$ tenemos

$$f = h_1 h_2 + v,$$

donde $v \in M[x]$ y $\mu(h_1) = \bar{g_1}$, $\mu(h_2) = g_2$. Ya que $\bar{g_1}$ y $\bar{g_2}$ son primos relativos si y sólo si h_1 y h_2 son primos relativos en $R[x]$, así existe λ_1 y λ_2 en $R[x]$, tal que

$$\lambda_1 h_1 + \lambda_2 h_2 = 1$$

Poniendo

$$h_{1,1} = h_1 + \lambda_2 v,$$
$$h_{2,1} = h_2 + \lambda_2 v$$

da

$$h_{1,1} h_{2,1} = f + \lambda_1 \lambda_2 v^2.$$

Por lo tanto $f \equiv h_{1,1} h_{2,1} (\text{mod } v^2)$., con $\mu(h_i, 1) = \mu(h_1)$, $i = 1, 2$ y $h_{1,1}, h_{2,1}$ primos relativos. En este punto podemos repetir el argumento aplicando esto a $h_{1,1}$ y $h_{2,1}$; Por iteración, podemos encontrar 2 polinomios $h_{1,t}$ y $h_{2,t}$ en $R[x]$, para todo entero positivo tal que

$$f \equiv h_{1,t_0} h_{2,t_0} \text{mod}(v^{2^t})$$

y

$$\mu(h_i, t) = \mu(h_i) \text{ con } i = 1, 2.$$

Obtenemos la afirmación (para el caso $n = 2$) eligiendo $g_i = h_i, t_0$, $1 \leq i \leq 2$. En general, si $\mu(f) = \bar{g_1} ... \bar{g_n}$, es suficiente observar que $\bar{g_1}$ es primo relativo a $\bar{g_i}$, $2 \leq i \leq n$, así $\{\bar{g_1}, ..., \bar{g_n}\}$ son primos relativos disjuntos. Poniendo $r = \bar{g_2} ... \bar{g_n}$ nos da $\mu(f) = \bar{g_1} r$ el cual completa la prueba. □

Bibliografía

[1] Adámek , J. *Foundations of Coding* ,Wiley, New York, 1991.

[2] Atiya, M.F. and Macdonald, I.G. *Introduction to Commutative Algebra* , Reading, Massachusetts- Menlo Park, California, Addison- Wesley Publishing Co., 1969.

[3] Bini, G. and Flamini, F. *Finite Commutative Rings and Their Applications*, Kluwer Academic Publisher, Springer. 2015.

[4] Friedberg, S. and Insel, A. *Algebra Lineal*, New Jersey, Pearson; Edición: 4th ed. 2002.

[5] Jacobson, N. *Basic Algebra I*, Second Edition, New York, W. H. Freeman and Co., 1995.

[6] Rincon, M. *Álgebra Lineal*, Second Edition, México D.F, W. H. Publidisa Mexicana, 2006.

[7] Lidl, Rudolf. and Niederreiter, H. *Finite Fields* ,New York, Cambridge Univerity Press, 1997.

[8] Ling, S. and Xing, C. *Coding Theory a First Course*, New York, Cambridge University Press; 1st edition (2004).

[9] Pless, V. *Introduction to the Theory of Error-Correcting Codes* ,Wiley, New York, 1982.

[10] Rotman, J. *Advanced Modern Algebra*, Prentice Hall; 1st edition (2002); 2nd printing (2003).

[11] Trappe, W. and Washington, L. *Introduction to Cryptography with Coding Theory*, New Jersey Pearson Prentice Hall; 1st edition (2002).