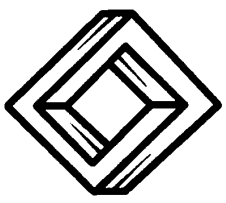
Publicaciones Electrónicas
Sociedad Matemática Mexicana

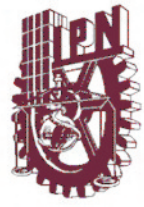
Álgebra Clásica
Segunda Edición

José Antonio Vargas Mendoza

www.sociedadmatematicamexicana.org.mx

Serie: Textos. Vol. 7 (2006)
ISBN: 968-9161-17-2





ALGEBRA CLASICA

Segunda edición

José Antonio Vargas Mendoza

Publicaciones Electrónicas
Sociedad Matemática Mexicana

ISBN 968-9161-17-2 en línea
ISBN 968-9161-18-0 en papel
ISBN 968-9161-19-9 en CD

Prólogo

Al iniciarse el siglo XXI, es cada vez mayor la cantidad y calidad de material que deben dominar los estudiantes de Matemáticas a nivel de licenciatura, independientemente de sus planes a futuro. Esto es particularmente cierto en lo que respecta al álgebra.

El presente libro intenta cubrir ese material agrupado en cinco capítulos correspondientes a grupos, anillos, teoría de Galois, álgebra lineal y temas complementarios. La situación ideal para quien aspire a adquirir una sólida base algebraica, para después completar estudios de postgrado, es que dedique un semestre a cada uno de los primeros cuatro capítulos del libro, de manera que tenga tiempo de adquirir el lenguaje, digerir los métodos y resultados aquí presentados; así como de interactuar con los problemas enunciados. También es posible diseñar cursos para un año de álgebra basándose en este libro; tal vez omitiendo algunas secciones.

El quinto capítulo puede ser usado como fuente para exposiciones de los alumnos o para lecturas adicionales.

Este es un proyecto ambicioso, que requiere de bastante trabajo tanto del alumno como del profesor. Por otro lado, cada vez hay más alumnos y profesores competentes capaces de cubrir el material aquí incluido.

El autor confiesa su mala intención de poner directamente en manos de alumnos destacados, ideas y retos que sus profesores de licenciatura tal vez no quieran darles.

El nombre del libro, Algebra Clásica, concuerda con el criterio usado para la elección de los temas a tratar y de su profundidad. El mayor prerequisito para su comprensión, es el interés por el tema, junto con un curso previo de álgebra lineal elemental.

Para esta segunda edición, se corrigieron múltiples errores, se agregaron ejercicios al capítulo 5 y se escribió una nueva demostración del Teorema de Frobenius que clasifica los anillos de división reales.

José Antonio Vargas M.
CIIDIR-Oaxaca, IPN
Oaxaca, Oax. México
Noviembre, 2009

Contenido

Capítulo 1
Grupos

1.1 Preliminares

En esta sección enunciamos ciertas propiedades de los enteros $\mathbb{Z}$ y de los enteros módulo n que se necesitarán inmediatamente.

Dados $a, b \in \mathbb{Z}$ con $b \neq 0$, existen $q, r \in \mathbb{Z}$ con $a = bq + r$ de manera que $0 \leq r < |b|$. Este es el **algoritmo euclideano**.

Un número entero $p > 1$ es **primo** cuando solamente es divisible por ± 1 y por $\pm p$.

Todo entero positivo distinto de 1 puede escribirse como producto de potencias positivas de primos. Esta expresión es única, en el sentido de que

$$n = p_1^{a_1} \cdots p_r^{a_r} = q_1^{b_1} \cdots q_s^{b_s},$$

donde $p_1, ..., p_r$ son primos distintos; y $q_1, ..., q_s$ también son primos distintos con $a_i, b_j > 0$ para todas i, j, implica que $r = s$ y para cada $1 \leq i \leq r$ existe $1 \leq j \leq s$ tal que j es único, $p_i = q_j$ y $a_i = b_j$.

Se dice que $c > 0$ es el **máximo común divisor** de m y n, escrito $c = \text{m.c.d.}\{m, n\}$ cuando c divide a m (escrito $c|m$), $c|n$ y $(d|m, d|n \Rightarrow d|c)$.

Cuando $m = p_1^{a_1} \cdots p_r^{a_r}$ y $n = p_1^{b_1} \cdots p_r^{b_r}$ con $a_i, b_j \geq 0$, entonces

$$\text{m.c.d.}\{m, n\} = p_1^{c_1} \cdots p_r^{c_r},$$

donde $c_i = \text{mín } \{a_i, b_i\}$, para todo i. Observamos que siempre es posible escribir dos números positivos m y n en esta forma, permitiendo que algunos exponentes sean cero.

Se dice que $s > 0$ es el **mínimo común múltiplo** de m y n, cuando $m|s$, $n|s$ y $(m|r, n|r \Rightarrow s|r)$. Esto se escribe así: $s = \text{m.c.m.}\{m, n\}$.

Cuando $m = p_1^{a_1} \cdots p_r^{a_r}$ y $n = p_1^{b_1} \cdots p_r^{b_r}$ con $a_i, b_j \geq 0$, entonces

$$\text{m.c.m.}\{m, n\} = p_1^{k_1} \cdots p_r^{k_r},$$

donde $k_i = \text{máx } \{a_i, b_i\}$, para todo i.

Por lo anterior, $(\text{m.c.d.}\{m, n\})(\text{m.c.m.}\{m, n\}) = mn$.

Dos enteros a y b son **primos relativos** cuando $\text{m.c.d.}\{a, b\} = 1$.

Los números naturales $\mathbb{N} = \{0, 1, 2, ...\}$ están **bien ordenados**; lo que quiere decir que satisfacen la siguiente condición:

Axioma 1.1 *Todo subconjunto no vacío de* $\mathbb{N}$ *tiene un elemento mínimo.*

Usando esta propiedad, tenemos la siguiente caracterización del m.c.d. de dos números:

Proposición 1.2 *El máximo común divisor de* m *y* n*, ambos no iguales a cero, es el mínimo elemento positivo del conjunto* $A = \{am+bn \mid a, b \in \mathbb{Z}\}$.

Demostración: A es claramente no vacío pues contiene a 0. Además, también contiene elementos positivos. Sea c el mínimo de ellos, de manera que existen $a, b \in \mathbb{Z}$ tales que $am + bn = c$.

Por el algoritmo euclideano, existen $q, r \in \mathbb{Z}$ tales que $m = cq + r$, satisfaciendo $0 \leq r < c$; pero $r = m - cq = (1 - qa)m + (-qb)n \in A$.

Siendo c mínimo positivo, se obtiene que $r = 0$, es decir, que $c|m$. Similarmente, $c|n$. Por último, $d|m, d|n \Rightarrow d|(am + bn) = c$. □

Pasamos ahora a definir los enteros módulo n:

Fijamos $0 < n \in \mathbb{Z}$ y definimos una relación $\sim$ en $\mathbb{Z}$ así:

$$a \sim b \Leftrightarrow n|(a - b).$$

Es inmediato que

- $\sim$ es **reflexiva**, esto es, que $a \sim a$ para toda $a \in \mathbb{Z}$;
- $\sim$ es **simétrica**: $a \sim b \Leftrightarrow b \sim a$;
- $\sim$ es **transitiva**: $a \sim b, b \sim c \Rightarrow a \sim c$.

Esto significa que $\sim$ es una **relación de equivalencia**, por lo que $\mathbb{Z}$ es la unión disjunta de las **clases de equivalencia**, es decir, de los conjuntos $\{m \in \mathbb{Z} \mid m \sim a\} = \{a + rn \mid r \in \mathbb{Z}\}$, que abreviamos así: $\overline{a}$.

Es común escribir $a \equiv b \pmod{n}$ cuando $a \sim b$.

Definimos operaciones en el conjunto de n elementos $\{\overline{0}, \overline{1}, ..., \overline{n-1}\}$ así:

$$\overline{a} + \overline{b} = \overline{a + b},\ \overline{a} \times \overline{b} = \overline{ab},$$

que están bien definidas como es fácil ver. En esta forma, este conjunto con esas operaciones son los **enteros módulo** n, escritos $\mathbb{Z}/n\mathbb{Z}$.

Ejercicios

1. Dados $a, b \in \mathbb{Z}$ con $b \neq 0$, demuestre que los números q, r tales que $a = bq + r$ con $0 \leq r < |b|$, cuya existencia garantiza el algoritmo euclideano, son únicos.

2. Sea T un subconjunto no vacío de $\mathbb{Z}$ tal que si $a, b \in T$, entonces $(a+b), (a-b) \in T$. Demuestre que T consiste de los múltiplos de algún entero m.

3. Demuestre que todo entero positivo se puede escribir de manera única como una suma de distintas potencias no negativas de 2.

1.2 Definiciones y Primeros Resultados

Un **grupo** G es un conjunto equipado con una operación (aquí escrita como multiplicación) tal que:

1. $a, b \in G \Rightarrow ab \in G$.

2. $a(bc) = (ab)c$ para todos $a, b, c \in G$.

   Esta propiedad se llama **asociatividad**.

3. Existe un elemento $1 \in G$ tal que $a1 = 1a = a$ para toda $a \in G$.

4. Para toda $a \in G$, existe $b \in G$ tal que $ab = ba = 1$.

Es inmediato que el elemento cuya existencia garantiza la condición 3 es único, pues si $1'$ satisface la condición 3, se tiene que

$$1 = 11' = 1'.$$

Este elemento es la **identidad** de G.

En vista de esta observación, la condición 4 tiene sentido; y además, dado a, se tiene que el elemento b de esa condición es único, pues si c también la satisface, entonces

$$ab = ac = 1 \Rightarrow b(ab) = b(ac) \Rightarrow (ba)b = (ba)c \Rightarrow b = c.$$

En esta situación, se escribe $b = a^{-1}$ y se dice que b es el **inverso** de a.

Es claro que $(a^{-1})^{-1} = a$ y que $(ab)^{-1} = b^{-1}a^{-1}$ para todos $a, b \in G$.

Cuando C es un conjunto finito, $\circ(C)$ denota el número de elementos de C y se llama el **orden** de C.

Se dice que un grupo G es **abeliano** cuando $ab = ba$ para todos $a, b \in G$.

Ejemplos. Como ejemplos de grupos tenemos los siguientes, para los que fijamos nuestra notación.

1. Los números enteros $\mathbb{Z}$ ante la suma.

2. Los números racionales $\mathbb{Q}$ ante la suma.

3. Los números reales $\mathbb{R}$ ante la suma.

4. Los números complejos $\mathbb{C}$ ante la suma.

5. Los enteros módulo n ante la suma, $\mathbb{Z}/n\mathbb{Z}$. Cuando eliminamos la multiplicación de este conjunto, escribimos Z_n y decimos que es el **grupo cíclico** de orden n.

6. Dado un conjunto arbitrario X, la colección de todas las biyecciones $f : X \to X$ forma un grupo ante la operación de composición de funciones, el **grupo simétrico** S_X. En caso de que $\circ(X) = n$, escribimos S_n. Los elementos de estos grupos se llam an **permutaciones**.

7. El conjunto de todas las matrices $n \times n$ con coeficientes en $\mathbb{Q}$, resp. en $\mathbb{R}$ ó en $\mathbb{C}$ y determinante $\neq 0$ forma un grupo ante la multiplicación de matrices, el **grupo general lineal** $GL_n(\mathbb{Q})$, resp. $GL_n(\mathbb{R})$ ó $GL_n(\mathbb{C})$.

8. El conjunto de todas las matrices $n \times n$ con coeficientes en $\mathbb{Q}$, resp. en $\mathbb{R}$ ó en $\mathbb{C}$ y determinante 1 forma un grupo ante la multiplicación de matrices, el **grupo especial lineal** $SL_n(\mathbb{Q})$, resp. $SL_n(\mathbb{R})$ ó $SL_n(\mathbb{C})$.

9. Por otra parte, los números naturales $\mathbb{N}$ no son un grupo ni ante la suma ni ante la multiplicación.

Dados un grupo G y un subconjunto $H \subseteq G$, se dice que H es un **subgrupo** de G cuando H es un grupo ante la misma operación de G. Esto lo escribimos así: $H < G$.

Ejemplo. $SL_n < GL_n$ sobre cualquier campo como $\mathbb{Q}, \mathbb{R}$ ó $\mathbb{C}$.

Para subconjuntos arbitrarios $A, B \subseteq G$, definimos los conjuntos

$$A^{-1} = \{a^{-1} \mid a \in A\} \text{ y } AB = \{ab \mid a \in A,\ b \in B\}.$$

Proposición 1.3 *Si $\varnothing \neq H \subseteq G$, entonces las siguientes condiciones son equivalentes:*
a) $H < G$.
b) $HH \subseteq H$ *y* $H^{-1} \subseteq H$.
c) $HH^{-1} \subseteq H$.

Demostración: *a)* $\Rightarrow$ *b)* y *b)* $\Rightarrow$ *c)* son claras. Veamos que *c)* $\Rightarrow$ *a)*:

Como existe $h \in H$, se tiene que $1 = hh^{-1} \in H$ y que por lo tanto $a \in H \Rightarrow a^{-1} = 1a^{-1} \in H$. Finalmente, tenemos que $a, b \in H \Rightarrow ab = a(b^{-1})^{-1} \in H$. $\square$

Observación. Si H es finito, entonces la condición $H^{-1} \subseteq H$ de *b)* es redundante, es decir, $HH \subseteq H \Rightarrow H^{-1} \subseteq H$, pues dado $h \in H$ se tiene que $hH \subseteq H$ y también $\circ(hH) = \circ(H) \Rightarrow hH = H$. Por tanto existe $k \in H$ con $hk = h$, así $k = 1$ y también existe $j \in H$ con $hj = 1$; pero $h^{-1} = j \in H$ en ese caso.

Teorema 1.4 (Lagrange) *Si G es un grupo finito y H es un subgrupo, entonces $\circ(H) \mid \circ(G)$.*

Demostración: Toda **clase lateral derecha** xH de H tiene $\circ(H)$ elementos. Si $xh_1 = yh_2 \in xH \cap yH$ con $h_1, h_2 \in H$, entonces $y^{-1}x = h_2h_1^{-1} \in H$, por lo que $xH = y(y^{-1}x)H = yH$. Se ve entonces que las clases laterales distintas son disjuntas. Si hay n de ellas, se tiene que $n(\circ(H)) = \circ(G)$. □

El **índice** de H en G, escrito $[G : H]$ es el número de clases laterales de H en G. Cuando G es finito, $[G : H] = \circ(G)/\circ(H)$, ver el Ejercicio 2, página 8.

Si $\{H_i\}$ es una colección de subgrupos de G para $i \in I$, entonces claramente $(\bigcap_{i\in I} H_i) < G$. Mientras que dado un subconjunto $A \subseteq G$, se tiene que la intersección de todas las H tales que $A \subseteq H < G$ es un subgrupo de G, escrito $\langle A\rangle$ y llamado el **subgrupo generado por** A.

Se dice que un grupo G es **cíclico** cuando existe un elemento $a \in G$ tal que $G = \langle a\rangle$. Es claro que Z_n es cíclico y de orden n. Usando la asociatividad, también es claro que todo grupo cíclico es abeliano.

Corolario 1.5 *Si G es un grupo finito de orden n, entonces $a^n = 1$ para toda $a \in G$.*

Demostración: Dado $1 \neq a \in G$, sea $H = \langle a\rangle$. Claramente se ve que $H = \{1, a, a^2, ..., a^{m-1}\}$, donde m es el mínimo entero positivo tal que $a^m = 1$. Así, $\circ(H) = m$, $m|n$ por el Teorema de Lagrange y $a^n = 1$. □

Definimos el **orden de un elemento** $a \in G$ como el orden de $\langle a\rangle$, escrito $\circ(a)$.

La **función** $\varphi : \mathbb{N} \to \mathbb{N}$ **de Euler** queda definida por $\varphi(1) = 1$ y para $n > 1$ por $\varphi(n) =$ número de enteros positivos menores que n y primos relativos a n. Por ejemplo, si p es primo, $\varphi(p^m) = p^m - p^{m-1}$.

Corolario 1.6 (Euler) *Si a y n son primos relativos, con n positivo, entonces $a^{\varphi(n)} \equiv 1 \pmod n$.*

Demostración: Observemos primero que si a y n son enteros primos relativos, entonces $a + kn$ y n también lo son, para todo $k \in \mathbb{Z}$. De manera que tiene sentido considerar a los elementos de $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, ..., \overline{n-1}\}$ que son primos con respecto a n. Este conjunto H forma un grupo multiplicativo de orden $\varphi(n)$, pues $HH \subseteq H$ y entonces $H^{-1} \subseteq H$ (vea la Observación previa). Como $\overline{a}$ pertenece a este grupo, se tiene que $a^{\varphi(n)} \equiv 1 \pmod n$. □

Corolario 1.7 (Fermat) *Si p es un número primo y a es un entero, entonces $a^p \equiv a \pmod p$.*

Demostración: Como $\varphi(p) = p-1$, se tiene que $a^{p-1} \equiv 1 \pmod p$ siempre que $p \nmid a$. En todo caso, $a^p \equiv a \pmod p$. □

Corolario 1.8 *Si $a > 1, n \geq 1$ son enteros, entonces $n \mid \varphi(a^n - 1)$.*

Demostración: Sea G el grupo multiplicativo de los enteros $(\mod a^n-1)$ primos con respecto a a^n-1. Entonces $\circ(G)=\varphi(a^n-1)$. Es claro que $\overline{a}\in G$ y que $\circ(a)=n$, por lo que $n \mid \varphi(a^n-1)$. □

Si A, B, C son conjuntos finitos, es fácil ver que

$$\circ(A\cup B)=\circ(A)+\circ(B)-\circ(A\cap B),$$

$$\circ(A\cup B\cup C)=\circ(A)+\circ(B)+\circ(C)-\circ(A\cap B)-\circ(B\cap C)-\circ(A\cap C)$$
$$+\circ(A\cap B\cap C).$$

Esto se generaliza como sigue.

Proposición 1.9 (Principio de inclusión y exclusión) *Si $A_1,...,A_n$ son conjuntos finitos, entonces*

$$\circ(\bigcup_i A_i)=\sum_i \circ(A_i)-\sum_{i<j}\circ(A_i\cap A_j)+\cdots+(-1)^{n-1}\circ(\bigcap_i A_i).$$

Demostración: Cada $a\in\cup_i A_i$ está contenido en exactamente un número $1\le t\le n$ de conjuntos A_i. Por eso, fijando t, el elemento a da origen en el lado derecho de nuestro enunciado a una contribución de

$$\binom{t}{1}-\binom{t}{2}+\cdots+(-1)^{t-1}\binom{t}{t}=\binom{t}{0}=1,$$

pues

$$\binom{t}{0}-\binom{t}{1}+\binom{t}{2}-\cdots+(-1)^t\binom{t}{t}=(1-1)^t=0.\ \square$$

Ahora aplicamos este resultado al cálculo de la función φ de Euler.

Corolario 1.10 *Si $p_1,...,p_m$ son los distintos primos que dividen a un entero positivo n, entonces*

$$\varphi(n)=n-\frac{n}{p_1}-\cdots-\frac{n}{p_m}+\frac{n}{p_1p_2}+\cdots+\frac{n}{p_{m-1}p_m}-\cdots=n\prod_{i=1}^{m}(1-\frac{1}{p_i}).$$

Demostración: $\varphi(n)$ enumera los elementos que quedan del conjunto $B=\{1,2,...,n\}$, al omitir aquellos que tienen un factor común no trivial con n.

Si A_i es el subconjunto de B formado por los números divisibles por p_i, tendremos que $\varphi(n)=\circ(B\setminus\cup_i A_i)$.

La conclusión se sigue de la Proposición 1.9 y de que

$$\circ(A_{i_1}\cap\cdots\cap A_{i_k})=\frac{n}{p_{i_1}\cdots p_{i_k}}.\ \square$$

Observación. Si m.c.d.$\{a,b\}=1$, entonces $\varphi(ab)=\varphi(a)\varphi(b)$.

Teorema 1.11 *Sea $G = \langle g \rangle$ un grupo cíclico de orden n. Entonces:*
a) Todo subgrupo de G es cíclico.
b) El orden de un elemento g^m es $n/\,\text{m.c.d.}\{m,n\}$.
c) El número de elementos $x \in G$ tales que $G = \langle x \rangle$ es $\varphi(n)$.
d) Para cada entero positivo r tal que $r|n$, existe un único subgrupo de orden r.

Demostración: *a)* Sea $H < G$. Definimos $T = \{i \in \mathbb{Z} \mid g^i \in H\}$. El conjunto T es cerrado ante la suma y la resta de sus elementos; así como ante la multiplicación por cualquier entero, por lo que existe $s \in \mathbb{N}$ tal que T consiste de los múltiplos de s. Así, $H = \langle g^s \rangle$.

b) El orden de un elemento g^m es el mínimo entero t tal que $n \mid mt$. Este número es $n/\,\text{m.c.d.}\{m,n\}$.

c) En vista de *b)*, un elemento g^i con $1 \leq i < n$ genera a G si y sólo si $\text{m.c.d.}\{n,i\} = 1$, es decir, si y sólo si i es primo con respecto a n. El número de posibilidades para i es entonces $\varphi(n)$.

d) También en vista de *b)*, un elemento g^i es de orden r si y sólo si $\text{m.c.d.}\{n,i\} = n/r$. De esta manera, $H = \langle g^{n/r} \rangle$ es de orden r y H contiene a todo elemento de G de orden r. □

Lema 1.12 *Si $H, K < G$, entonces las siguientes condiciones son equivalentes:*
a) $HK < G$.
b) $HK \subseteq KH$.
c) $HK = KH$.

Demostración: *c)* $\Rightarrow$ *b)* es claro.

Veamos que *b)* $\Rightarrow$ *a)* verificando que se satisface la condición *c)* de la Proposición 1.3 para HK:

$$(HK)(HK)^{-1} = H(KK^{-1})H^{-1} \subseteq HKH^{-1} = HK^{-1}H^{-1};$$

pero $HK \subseteq KH \Rightarrow K^{-1}H^{-1} \subseteq H^{-1}K^{-1}$ de manera que

$$HK^{-1}H^{-1} \subseteq HH^{-1}K^{-1} \subseteq HK^{-1} = HK.$$

Así podemos concluir que $(HK)(HK)^{-1} \subseteq HK$ y que $HK < G$.

Finalmente, *a)* $\Rightarrow$ *c)*, porque $HK = (HK)^{-1} = K^{-1}H^{-1} = KH$. □

Corolario 1.13 *Si G es un grupo abeliano y $H, K < G$, entonces también $HK < G$.*

Lema 1.14 *Si H, K son subgrupos finitos de G, entonces*

$$\circ(HK) = \frac{\circ(H) \circ (K)}{\circ(H \cap K)}.$$

Demostración: Esto es consecuencia de observar que $h_1k_1 = h_2k_2 \Leftrightarrow h_2^{-1}h_1 = k_2k_1^{-1} \in H \cap K$, para todas $h_i \in H, k_i \in K$. □

Proposición 1.15 (Poincaré) *Si H, K son subgrupos de índice finito en un grupo G, entonces $H \cap K$ también tiene índice finito.*

Demostración: Para cualquier $x \in G$ arbitrario, es claro que se tiene $x(H \cap K) = (xH) \cap (xK)$, por lo que solamente hay un número finito de posibilidades para clases laterales de $(H \cap K)$. □

Ejercicios

1. Un **monoide** es un conjunto con una operación que satisface las primeras tres condiciones para ser grupo. Dé 3 ejemplos de monoides que no sean grupos.

2. Sean G un grupo y H un subgrupo. Construya una biyección del conjunto de las clases laterales izquierdas de H en G al conjunto de las clases laterales derechas de H en G.

3. Sean H y K subgrupos finitos de un grupo G. Demuestre que el número de elementos de una **clase lateral doble**, definida como $HxK = \{hxk \mid h \in H, k \in K\}$ es igual a $\circ(H)[K : (x^{-1}Hx \cap K)]$.

4. Sean $m, n \in \mathbb{N}$ con $r = \text{m.c.d.}\{m, n\}$. Demuestre que
$$\varphi(mn) = \varphi(m)\varphi(n)\frac{r}{\varphi(r)}.$$

5. Para cierto entero positivo h, se tiene que el número $2^h + 1 = p$ es primo.

   a) Demuestre que el orden de 2 en el grupo multiplicativo de $\mathbb{Z}/p\mathbb{Z}$ es $2h$.

   b) Demuestre que $2h \mid (p-1) = 2^h$.

   c) Demuestre que h es una potencia de 2.

6. *a*) Sean d y n enteros positivos tales que $d|n$. Demuestre que el número de enteros i tales que $0 < i \leq n$ y $\text{m.c.d.}\{i, n\} = d$, es $\varphi(n/d)$.

   b) Demuestre que $\sum_{d|n} \varphi(n/d) = n$.

7. Sean A y B conjuntos finitos con $\circ(A) = m$, $\circ(B) = n$ y $m \geq n$.

   a) Calcule el número de funciones $f : A \to B$.

   b) Demuestre que el número de funciones suprayectivas $f : A \to B$ es
$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^m.$$

1.3 Subgrupos Normales

Se dice que un subgrupo N de G es **normal** cuando $xNx^{-1} \subseteq N$ para toda $x \in G$. Esto es claramente equivalente a $xNx^{-1} = N$ para toda $x \in G$ y se escribe $N \lhd G$.

Proposición 1.16 *Para $N < G$, las siguientes condiciones son equivalentes:*

a) $N \lhd G$.

b) Toda clase lateral izquierda de N es una clase lateral derecha de N (o recíprocamente). Más precisamente, $Nx = xN$ para toda $x \in G$.

c) El producto de dos clases laterales izquierdas (resp. derechas) de N es una clase lateral izquierda (resp. derecha) de N.

Demostración: *a)* $\Rightarrow$ *b)*: Se tiene que $xNx^{-1} = N$ para toda $x \in G$. Por tanto, $xN = Nx$ para toda $x \in G$.

b) $\Rightarrow$ *c)*: $(Na)(Nb) = N(aN)b = N(Na)b = Nab$ para todas $a, b \in G$.

c) $\Rightarrow$ *a)*: Para toda $x \in G$, $NxNx^{-1}$ es una clase lateral izquierda de N que contiene a 1; por tanto $NxNx^{-1} = N$ y así $xNx^{-1} \subseteq N$. □

Observaciones. Las siguientes afirmaciones son todas fáciles de verificar.

1. Todo subgrupo de índice 2 es normal. Esto es consecuencia de la equivalencia a) $\Leftrightarrow$ b) de la Proposición 1.16.

2. Si G es abeliano, entonces todo subgrupo de G es normal.

3. Toda intersección de subgrupos normales es un subgrupo normal.

4. Si $N \lhd G$ y $H < G$, entonces $(N \cap H) \lhd H$.

5. Si $N \lhd G$ y $H < G$, entonces $NH < G$ y $N \lhd NH$.

Ejemplo. Sea $\mathsf{H} = \{\pm 1, \pm i, \pm j, \pm k\}$ el conjunto de 8 elementos con multiplicación dada por $i^2 = j^2 = k^2 = -1, ij = k = -ji, jk = i = -kj, ki = j = -ik, (-1)^2 = 1, (-1)(\pm i) = \mp i, (-1)(\pm j) = \mp j, (-1)(\pm k) = \mp k$; este es el **grupo de cuaternios**. Todos los subgrupos de H son normales y H no es abeliano, por lo que el recíproco de la Observación 2 es falso.

Teorema 1.17 *Si $N \lhd G$, entonces el conjunto G/N de las clases laterales izquierdas de N en G es un grupo.*

Demostración: Esto es claro porque $(Na)(Nb) = Nab$ y $(Na)^{-1} = Na^{-1}$ para todas $a, b \in G$, ante la multiplicación de bloques, que es asociativa con identidad N. □

Se dice que un grupo G es **simple** cuando sus únicos subgrupos normales son $\{1\}$ y G.

Para $a, b \in G$ definimos su **conmutador** como $(a, b) = aba^{-1}b^{-1}$. Como a y b conmutan si y sólo si $(a, b) = 1$, se puede interpretar al conmutador de ellos como una medida de su falta de conmutatividad.

Definimos al **grupo derivado** G' de un grupo G como

$$G' = \langle (a, b) \mid a, b \in G \rangle.$$

Proposición 1.18 *Si $N \lhd G$, entonces G/N* es abeliano $\Leftrightarrow G' \subseteq N$.

Demostración: $Nab = Nba \Leftrightarrow Naba^{-1}b^{-1} = N \Leftrightarrow aba^{-1}b^{-1} \in N$. □

Proposición 1.19 *Si $H < G$ y $G' \subseteq H$, entonces $H \lhd G$.*

Demostración: Dados $h \in H$ y $g \in G$ arbitrarios, $ghg^{-1}h^{-1} \in G' \subseteq H$; por consiguiente $ghg^{-1} \in H$. □

Proposición 1.20 *Si $M, N \lhd G$ y $M \cap N = \{1\}$, entonces $mn = nm$ para todos $m \in M$ y $n \in N$.*

Demostración: $(m, n) = (mnm^{-1})n^{-1} = m(nm^{-1}n^{-1}) \in M \cap N$. Por tanto, $(m, n) = 1$. □

Ejercicios

1. Sean G un grupo finito y $H < G$. Demuestre que (H es normal) $\Leftrightarrow$ (todas las clases laterales dobles HaH tienen el mismo número de elementos).

2. Para $a, b \in \mathbb{R}$, sea $T_{a,b} : \mathbb{R} \to \mathbb{R}$ la transformación dada por $T_{a,b}(x) = ax + b$ para todo $x \in \mathbb{R}$.

   a) Demuestre que $G = \{T_{a,b} \mid a \neq 0\}$ es un grupo ante la composición de funciones.

   b) Demuestre que $U = \{T_{1,b}\}$ es un subgrupo normal de G.

3. Sean G un grupo abeliano infinito y T la colección de todos los elementos de G de orden finito. Demuestre que $T \lhd G$.

1.4 Morfismos

Un **homomorfismo** o simplemente un **morfismo** de grupos es una función $f : G \to H$, donde G y H son grupos, tal que $f(ab) = f(a)f(b)$ para todos $a, b \in G$.

Ejemplos. Las siguientes funciones son morfismos de grupos:

1. Para todo grupo G, la función identidad $id : G \to G$, donde $id(g) = g$ para todo $g \in G$.

2. Para todo grupo G, la función constante $1 : G \to \{1\}$, donde $1(g) = 1$ para todo $g \in G$.

3. El determinante $\det : GL_n(\mathbb{Q}) \to \mathbb{Q}^\star$, donde $\mathbb{Q}^\star$ es el grupo multiplicativo de los elementos distintos de cero de $\mathbb{Q}$.

4. Si $N \lhd G$, entonces $\varphi : G \to G/N$ dado por $\varphi(g) = Ng$ es un morfismo llamado **natural**. Este morfismo es suprayectivo.

El **núcleo** de un morfismo $f : G \to H$ es el conjunto $\{x \in G \mid f(x) = 1\}$, se escribe $\ker f$. Un **isomorfismo** es un morfismo $\varphi : G \to H$ tal que admite un **morfismo inverso**, esto es, $\psi : H \to G$ de manera que $\psi \circ \varphi = id_G$ y también $\varphi \circ \psi = id_H$.

Dos grupos G y H son **isomorfos** cuando existe un isomorfismo entre ellos $f : G \to H$. Esto se escribe $G \cong H$.

El núcleo de "det" del ejemplo 3 es $SL_n(\mathbb{Q})$, mientras que el núcleo del ejemplo 4 es N.

Observaciones. Las siguientes afirmaciones para un morfismo $f : G \to H$ son fáciles de verificar:

1. $f(1) = 1$ y $f(x^{-1}) = f(x)^{-1}$ para todo $x \in G$.

2. $\operatorname{Im} f < H$.

3. $\ker f \lhd G$.

4. Si $f(a) = b$, entonces $f^{-1}(b) = Ka$, donde $K = \ker f$.

5. f es un isomorfismo $\Leftrightarrow$ f es suprayectivo y $\ker f = \{1\}$.

Como consecuencia de 2 y 3, tenemos que dados un grupo G y un subgrupo normal N, entonces existen un grupo E y un morfismo suprayectivo $f : G \to E$ con núcleo N.

Teorema 1.21 *Sea $f : G \to H$ un morfismo suprayectivo con núcleo K. Entonces $K \lhd G$ y $G/K \cong H$.*

Demostración: Sea $\varphi : G/K \to H$ la función dada por $\varphi(Kx) = f(x)$. Es fácil ver que φ está bien definida y que es un isomorfismo. □

Teorema 1.22 *Sean $H < G$ y $N \lhd G$. Entonces*

$$(H \cap N) \lhd H \text{ y } H/(H \cap N) \cong (HN/N).$$

Demostración: Sean $f : G \to G/N$ el morfismo natural y g la restricción de f a H. Observamos que $\operatorname{Im} g = HN/N$ y que $\ker g = H \cap N$. La conclusión es consecuencia del teorema anterior. □

Teorema 1.23 *Sean $N \subseteq H \subseteq G$ tres grupos con N y H normales en G. Entonces $(H/N) \lhd (G/N)$ y $(G/N)/(H/N) \cong (G/H)$.*

Demostración: Sea $f : G/N \to G/H$ la función dada por $f(Na) = Ha$ para toda $a \in G$. Es fácil verificar que f es un morfismo suprayectivo con núcleo H/N. Para terminar aplicamos el Teorema 1.21. □

Teorema 1.24 *Sean G y E grupos y sea $f : G \to E$ un morfismo suprayectivo con núcleo N. La acción de f en subconjuntos de G da origen a una biyección de $\{$los subgrupos de G que contienen a $N\}$ a $\{$los subgrupos de $E\}$. Esta biyección preserva inclusiones y normalidad. Además, si $N \subseteq A < B < G$ y $A' = f(A)$, $B' = f(B)$, entonces $[B : A] = [B' : A']$ y $A \lhd B \Leftrightarrow A' \lhd B'$, en cuyo caso $B/A \cong B'/A'$.*

Demostración: Si $H < G$ es tal que $N \subseteq H$, entonces $f(H) < E$, por lo que se puede definir $\varphi : \{$subgrupos de G que contienen a $N\} \to \{$subgrupos de $E\}$ así: $\varphi(H) = \{f(h) \mid h \in H\}$; también es claro que φ preserva inclusiones y que es suprayectiva: $T < E \Rightarrow f^{-1}(T) < G$ con $N \subseteq f^{-1}(T)$; y como f es suprayectiva, $ff^{-1}(T) = T$.

Veamos que φ es una función inyectiva: Sean $H_1, H_2 < G$ tales que $N \subseteq H_1, H_2$ y $\varphi(H_1) = \varphi(H_2)$. Si $a \in H_1$, entonces existe $b \in H_2$ tal que $\varphi(a) = \varphi(b)$; por tanto $ab^{-1} \in N \subseteq H_2$ y así $a = (ab^{-1})b \in H_2 b = H_2$. Esto demuestra que $H_1 \subseteq H_2$. La otra inclusión se obtiene de manera análoga.

Es claro que si se tiene que $N \subseteq A < B < G$ y que $A \lhd B$, entonces $A' \lhd B'$. Recíprocamente, si $A' \lhd B'$ y $b \in B$, entonces $bAb^{-1} \subseteq AN \subseteq A$. Además, $B'/A' \cong (B/N)/(A/N) \cong B/A$, por el Teorema 1.23.

Finalmente, $N \subseteq A < B < G \Rightarrow [B : A] = [B' : A']$, porque f envía las clases laterales de A en B a las clases laterales de A' en B' de manera biyectiva. □

Ejercicios

1. Sea φ un morfismo de grupos. Demuestre que φ es un isomorfismo si y sólo si φ es biyectivo.

2. Sea $\varphi : G \to H$ un morfismo de grupos finitos con núcleo K. Sean $A < B$ subgrupos de G y $A' < B'$ sus imágenes. Demuestre que

$$[B : A] = [B' : A'][(B \cap K) : (A \cap K)].$$

1.5 Conjugación y Automorfismos

Un **endomorfismo** de grupos es un morfismo $f : G \to G$ de un grupo G en sí mismo. Un **automorfismo** de G es un isomorfismo de G en G. El conjunto de los automorfismos de un grupo G forma un grupo ante la composición de funciones; que se escribe $\operatorname{Aut} G$.

Dados un grupo G y un elemento arbitrario $a \in G$, definimos una función $i_a : G \to G$, llamada **conjugación** con a, así: $i_a(x) = axa^{-1}$. Como $i_a(xy) = i_a(x)i_a(y)$ y también $i_a(x^{-1}) = [i_a(x)]^{-1}$, tenemos que $i_a \in \operatorname{Aut} G$.

Definimos $\operatorname{Int} G = \{i_a | a \in G\}$, el conjunto de los **automorfismos internos** de G.

Como $i_a i_b(x) = (ab)x(ab)^{-1} = i_{ab}(x)$ es cierto para toda $x \in G$, tenemos que $i_a \circ i_b = i_{ab}$. Similarmente, $i_{a^{-1}} = i_a^{-1}$, de manera que $\operatorname{Int} G < \operatorname{Aut} G$.

Proposición 1.25 *$\operatorname{Int} G \lhd \operatorname{Aut} G$.*

Demostración: Sean $\alpha \in \operatorname{Aut} G$, $x \in G$ arbitrarios, entonces es válido que $(\alpha i_x \alpha^{-1})(g) = \alpha(x\alpha^{-1}(g)x^{-1}) = \alpha(x)g\alpha(x)^{-1}$ para toda $g \in G$, por lo que $\alpha i_x \alpha^{-1} = i_{\alpha(x)}$, que demuestra el enunciado. □

Se dice que dos elementos a y b de un grupo G son **conjugados** en G cuando existe $x \in G$ tal que $b = xax^{-1}$. Evidentemente, conjugación es una relación de equivalencia; sus clases de equivalencia se llaman **clases de conjugación**.

Se dice que un subgrupo H de un grupo G es **característico** cuando $f(H) \subseteq H$ para toda $f \in \operatorname{Aut} G$.

Dados un grupo G y un subconjunto $A \subseteq G$, se definen el **centralizador** de A en G, escrito $Z_G(A)$, como $\{x \in G \mid xa = ax \text{ para toda } a \in A\}$ y el **normalizador** de A en G, escrito $N_G(A)$, como $\{x \in G \mid xA = Ax\}$.

El subgrupo $Z_G(G)$ se llama el **centro** de G y se escribe Z.

Observaciones. Las siguientes afirmaciones son fáciles de demostrar:

1. Todo subgrupo característico es normal.
2. Todo subgrupo normal es una unión de clases de conjugación.
3. $Z_G(A), N_G(A) < G$ para cualquier subconjunto $A \subseteq G$.
4. $Z_G(A) \lhd N_G(A)$ para cualquier subconjunto $A \subseteq G$.
5. Z es un subgrupo característico de G.

Teorema 1.26 *$G/Z \cong \operatorname{Int} G$.*

Demostración: Sea $f : G \to \operatorname{Int} G$ la función dada por $f(a) = i_a$ para toda $a \in G$. Este es un morfismo suprayectivo con núcleo Z. □

Ejercicios

1. Sea G un grupo con $Z = \{1\}$. Demuestre que $Z_{\operatorname{Aut} G}(\operatorname{Int} G) = \{1\}$.
2. Sea Z el centro de G. Demuestre que G/Z cíclico $\Rightarrow G$ abeliano.
3. Sea G un grupo finito con $f \in \operatorname{Aut} G$ tal que
   - $f^2 = 1$,
   - $f(x) = x \Rightarrow x = 1$.

   Demuestre que G es abeliano.
4. Describa $\operatorname{Aut} Z_n$, donde n es un entero positivo.
5. Sea G un grupo con exactamente un elemento a de orden 2. Demuestre que a es central.
6. Sea G un grupo con exactamente dos clases de conjugación, y que contiene un elemento de orden $n > 1$. Demuestre que $\circ(G) = 2$.
7. Sean G un grupo finito y $N \lhd G$ tal que $\circ(N) = n$, $[G : N] = m$ con $\text{m.c.d.}\{m, n\} = 1$. Demuestre que N es un subgrupo característico.

1.6 Acciones de Grupos

Se dice que un grupo G **actúa** en un conjunto X cuando se tiene un morfismo $f : G \to S_X$. Esto es equivalente a tener una función

$$\varphi : G \times X \to X,$$

tal que al escribir $\varphi(g, x) = g \cdot x$, se cumplan las condiciones:

- $(gh) \cdot (x) = g \cdot (h \cdot x)$, para todos $g, h \in G, x \in X$.
- $1 \cdot x = x$ para toda $x \in X$.

En estas condiciones, definimos los siguientes conceptos: Para cada elemento $x \in X$, el **estabilizador** de x es $G_x = \{g \in G \mid g \cdot x = x\}$, que es claramente un subgrupo de G. La **órbita** de x es $G \cdot x = \{g \cdot x \mid g \in G\}$.

Una acción es **transitiva** cuando el número de órbitas es 1, es decir, cuando existe $x \in X$ tal que $G \cdot x = X$. Se dice que $x \in X$ es un **punto fijo** cuando $G_x = G$. El conjunto de los puntos fijos de X se escribe X^G.

Un ejemplo de acción del grupo G sobre el conjunto G es conjugación, donde $f(g) = i_g$ para toda $g \in G$. Aquí las órbitas son las clases de conjugación, el estabilizador de un elemento x es su centralizador $Z(x)$; y un elemento es un punto fijo cuando pertenece al centro del grupo.

Teorema 1.27 *Si el grupo G actúa en el conjunto X, entonces X es la unión disjunta de las órbitas. Existe una biyección de $\{$los elementos de la órbita de $x\}$ a $\{$las clases laterales de G_x en $G\}$. En particular, $\circ(G \cdot x) = [G : G_x]$.*

Demostración: La condición "a, b pertenecen a una órbita" define una relación de equivalencia, por lo que la primera afirmación es clara.

Definimos $\varphi : G/G_x \to G \cdot x$ así: $\varphi(aG_x) = a \cdot x$. Claramente φ es una función suprayectiva, $\varphi(aG_x) = \varphi(bG_x) \Rightarrow b^{-1}a \in G_x \Rightarrow bG_x = b(b^{-1}a)G_x = aG_x$, por lo que φ es una biyección. La tercera afirmación es inmediata. □

En el caso de conjugación, escribimos la órbita de x como $C(x)$, esta es la clase de conjugación de x. Aquí tenemos $\circ(C(x)) = [G : Z(x)]$; y para grupos finitos, la **ecuación de clase**

$$\circ(G) = \sum_{C(a)} \frac{\circ(G)}{\circ(Z(a))}, \tag{1.1}$$

donde la suma se toma sobre las distintas clases de conjugación de G.

Teorema 1.28 *Si G es un grupo, $\circ(G) = p^n$ con p un número primo y $n \geq 1$, entonces $Z \neq \{1\}$.*

Demostración: En la ecuación de clase

$$\circ(G) = \circ(Z) + \sum_{C(a) \neq \{a\}} \frac{\circ(G)}{\circ(Z(a))},$$

donde la suma se toma sobre clases de conjugación de elementos no centrales, se tiene que

$$p \mid \frac{\circ(G)}{\circ(Z(a))},$$

siempre, por lo que $p \mid \circ(Z)$ y así $\circ(Z) > 1$. □

Corolario 1.29 *Si G es un grupo y $\circ(G) = p^2$ con p un número primo, entonces G es abeliano.*

Demostración: Como $\circ(Z) = p$ ó p^2, es suficiente ver que $\circ(Z) \neq p$.

Supongamos que $\circ(Z) = p$ y que $a \notin Z$, entonces $Z \subset Z(a) < G$.

Por el Teorema de Lagrange, $Z(a) = G$ y entonces $a \in Z$, que es una contradicción. □

El siguiente teorema garantiza que dado un grupo G, siempre existen un conjunto X y una acción inyectiva $f : G \to S_X$.

Teorema 1.30 (Cayley) *Todo grupo G es isomorfo a un grupo de permutaciones.*

Demostración: Definimos una función $f : G \to S_G$ así: $f(a) = f_a$ es multiplicación izquierda por a, es decir, $f_a(x) = ax$ para toda $x \in G$.

De esta manera, $f_a \in S_G$ y también $(f_a \circ f_b)(x) = a(bx) = abx = f_{ab}(x)$ para toda $x \in G$, por lo que $f_a \circ f_b = f_{ab}$ y entonces f es un morfismo, cuyo núcleo $\{a \in G \mid ax = x \text{ para toda } x \in G\}$ es $\{1\}$. La conclusión es que $G \cong \operatorname{Im} f$. □

De esta manera, vemos que cualquier grupo está contenido en un grupo simétrico. El siguiente resultado es una generalización.

Teorema 1.31 *Sean G un grupo, H un subgrupo y $A = G/H$ el conjunto de las clases laterales derechas de H en G. Entonces existe una acción $f : G \to S_A$ cuyo núcleo es el máximo subgrupo normal de G contenido en H.*

Demostración: Definimos una función $f : G \to S_A$ como en el teorema anterior: $f(a) = f_a$ es multiplicación izquierda por a, es decir, $f_a(xH) = axH$ para toda $x \in G$.

Claramente, f es un morfismo. Sea K su núcleo.

Como $k \in K \Rightarrow kH = H$, se ve que $k \in H$ y que $K \subseteq H$.

Sabemos que $K \lhd G$. Sea $N \lhd G$ tal que $N \subseteq H$.

Entonces, $n \in N \Rightarrow g^{-1}ng \in N \subseteq H$ para todas $g \in G$ y $n \in N$; así $g^{-1}ngH = H$, es decir, $ngH = gH$ para todas $g \in G$ y $n \in N$. La conclusión es que $N \subseteq K$. □

Corolario 1.32 *Si G es un grupo y H es un subgrupo de G distinto de G tal que $\circ(G) \nmid [G : H]\,!$, entonces H contiene un subgrupo normal no trivial de G. En particular, G no es simple.*

Demostración: Sea $f : G \to S_A$ como en el teorema anterior, con núcleo K. Sabemos que $K \lhd G$ y que $K \subseteq H$. Si $K = \{1\}$, entonces f es inyectivo y $\circ(G) = \circ(\operatorname{Im} f) \mid [G : H]\,!$ □

Teorema 1.33 (Cauchy-Frobenius) *Si p es un número primo y $p \mid \circ(G)$, entonces el número de soluciones en G de la ecuación $x^p = 1$ es un múltiplo de p. En particular, existe al menos un elemento de G de orden p.*

Demostración: Sea $A = \{(x_1, ..., x_p) \in G^p \mid x_1 \cdots x_p = 1\}$.

Entonces $\circ(A) = \circ(G)^{p-1}$, pues $x_1, ..., x_{p-1}$ pueden elegirse entre los elementos de G arbitrariamente, mientras que $x_p = (x_1 \cdots x_{p-1})^{-1}$; por tanto, $p \mid \circ(A)$.

El conjunto A es la unión disjunta de las clases de equivalencia

$$(x_1, ..., x_p) \sim (x_2, ..., x_p, x_1) \sim \cdots \sim (x_p, x_1, ..., x_{p-1}),$$

provenientes de la relación donde son equivalentes las colecciones ordenadas de p elementos de G que se obtienen por rotación.

Cada una de estas clases de equivalencia tiene o bien un único elemento en el caso de que $x_1 = x_2 = \cdots = x_p$, ó bien p elementos en cualquier

otro caso. Esto es debido a que podemos considerar nuestros índices como elementos de $\mathbb{Z}/p\mathbb{Z}$, donde la igualdad de dos elementos de una clase de equivalencia implica

$$x_i = x_{i+j} = x_{i+2j} = \cdots,$$

para todo i con $1 \leq i \leq p$; y algún j con $1 \leq j \leq p-1$. Así tenemos que $x_1 = x_2 = \cdots = x_p$, pues

$$\mathbb{Z}/p\mathbb{Z} = \{i, i+j, i+2j, ...\}.$$

De lo anterior se obtiene que el número de elementos $x \in G$ tales que $x^p = 1$ es un múltiplo de p y como $x = 1$ es uno de ellos, existe al menos otro, es decir, un elemento de orden p. □

Este teorema es un recíproco parcial del Teorema de Lagrange, al afirmar que existe un subgrupo H de orden p del grupo G suponiendo que $p \mid \circ(G)$ y que p es primo.

Ejercicios

1. Sea G un grupo de orden impar. Demuestre que el número de clases de conjugación es impar.

2. El grupo G actúa transitivamente en el conjunto no vacío X.

   a) Demuestre que los estabilizadores de los distintos puntos de X son conjugados.

   b) Sea $\text{Tran}(x, y) = \{g \in G \mid g \cdot x = y\}$. Demuestre que $\text{Tran}(x, y)$ es una clase lateral de G_x.

3. Sean H y K subgrupos de G. Demuestre que el número de conjugados de H con elementos de K es $[K : N_G(H) \cap K]$.

4. Sean G un grupo infinito y H un subgrupo propio de índice finito. Demuestre que G contiene un subgrupo normal propio de índice finito.

5. Sean G un grupo finito y H un subgrupo de índice 2. Demuestre que el número de conjugados de $x \in H$ en H es n ó $n/2$, si n es el número de conjugados de x en G.

6. Demuestre el siguiente **Teorema de Burnside:** Si un grupo finito G actúa en un conjunto finito X, entonces el número de órbitas es

   $$\frac{1}{\circ(G)} \sum_{g \in G} \circ(X^g),$$

   donde $X^g = \{x \in X \mid g \cdot x = x\}$ para cada $g \in G$.

1.7 El Grupo Simétrico

Sabemos que para todo conjunto X, las biyecciones de X forman el grupo llamado simétrico S_X. Cuando X es finito, con n elementos, escribimos S_n e identificamos $X = \{1, 2, ..., n\}$.

Es un ejercicio fácil demostrar que $\circ(S_n) = n!$.

Si $\sigma \in S_n$, entonces el grupo $\langle\sigma\rangle$ actúa naturalmente en $\{1, 2, ..., n\}$ y descompone a este conjunto en órbitas que son también llamadas órbitas de σ.

Una notación eficiente para las permutaciones consiste en escribir una tras otra las órbitas de cada permutación, como

$$\sigma = (1, \sigma(1), \sigma^2(1), ...) \cdots$$

Así por ejemplo, $\sigma \in S_3$ tal que $\sigma(1) = 3, \sigma(3) = 2, \sigma(2) = 1$ se escribe $\sigma = (132)$. Cada órbita así escrita se llama **ciclo** y por ejemplo (132) es un 3-ciclo. Los puntos fijos no se escriben.

Otro ejemplo es $\tau \in S_4$ tal que $\tau(1) = 3, \tau(2) = 4, \tau(3) = 1, \tau(4) = 2$. Aquí, $\tau = (13)(24)$.

Estos ejemplos se generalizan así:

Proposición 1.34 *Toda permutación es un producto de ciclos disjuntos: sus órbitas.*

Observaciones. Es conveniente mencionar que:

1. El producto de permutaciones $\alpha\beta$ representa la biyección que resulta de aplicar primero β y después α. Esta es la composición de funciones normalmente escrita $\alpha \circ \beta$.

2. En la Proposición 1.34, las órbitas de una permutación dada son únicas; la escritura de un ciclo no es única, pues depende del número inicial. Se tiene unicidad en la escritura de un ciclo si exigimos que su número inicial sea mínimo.

3. Los ciclos disjuntos conmutan entre sí.

4. Los ciclos de una permutación admiten un orden total de acuerdo con sus elementos mínimos.

5. Se tiene unicidad en las expresiones de la Proposición 1.34 si exigimos que sus ciclos se escriban en orden (interno y externo).

Consideremos al conjunto $\mathbb{Q}[X_1, ..., X_n]$ de los polinomoios en n variables con coeficientes en $\mathbb{Q}$. El grupo S_n actúa en este conjunto de manera natural al decretar para todas $a \in \mathbb{Q}; 1 \leq i \leq n; \sigma \in S_n; f, g \in \mathbb{Q}[X_1, ..., X_n]$ que:

1. $\sigma(a) = a$.

2. $\sigma(X_i) = X_{\sigma(i)}$.

3. $\sigma(f+g) = \sigma(f) + \sigma(g)$.

4. $\sigma(fg) = \sigma(f)\sigma(g)$.

El polinomio $h = \Pi_{i<j}(X_i - X_j)$ tiene la propiedad especial de que ante esta acción, $\sigma(h) = \pm h$, por lo que si $n \geq 2$, la órbita de h tiene 2 elementos $\{\pm h\}$, ya que $(12)h = -h$.

Sea A_n el estabilizador de h. Entonces $A_n \lhd S_n$, por ser de índice 2. El grupo A_n se llama **alternante** y sus elementos **permutaciones pares**.

La acción de S_n en $\{\pm h\}$ da origen a un morfismo suprayectivo de grupos, llamado **signo**, sgn : $S_n \to \{\pm 1\}$, cuyo núcleo es A_n. Un poco más generalmente, si tenemos $H < S_n$ tal que $H \not\subseteq A_n$, "sgn" se restringe a H y sigue siendo suprayectivo, por lo que también el núcleo de la restricción es de índice 2. En otras palabras, $[H : (H \cap A_n)] = 2$, es decir, que en cualquier grupo de permutaciones tal que no todos sus elementos son pares, exactamente la mitad lo son.

Proposición 1.35 *Las clases de conjugación de S_n son los conjuntos de permutaciones con la misma descomposición cíclica. El número de clases de conjugación de S_n es el número de particiones de n, es decir, el número de maneras en que n se puede escribir como $n = n_1 + \cdots + n_m$, con cada n_i un entero positivo.*

Demostración: Todo es consecuencia de la observación de que si $\alpha = (i_1 i_2 \cdots i_m) \cdots$ es la descomposición en ciclos disjuntos de α, entonces para toda $\sigma \in S_n$ se tiene que

$$\sigma\alpha\sigma^{-1} = (\sigma(i_1)\sigma(i_2)\cdots\sigma(i_m))\cdots \quad \square$$

Las **transposiciones** son los conjugados de (12).

Proposición 1.36 *Toda permutación es un producto de transposiciones, es decir, las transposiciones generan S_n. Mejor aún, las transposiciones de la forma $(1a)$ con a arbitrario generan S_n.*

Demostración: Es suficiente observar que son válidas las identidades $(i_1 i_2 \cdots i_m) = (i_1 i_m) \cdots (i_1 i_3)(i_1 i_2)$ y $(ab) = (1b)(1a)(1b)$, si $a, b \neq 1$. $\square$
Observaciones. Las siguientes afirmaciones son inmediatas:

1. El signo de toda transposición es -1.

2. El signo de un m-ciclo es $(-1)^{m-1}$. (Ver la última demostración).

3. El orden de un m-ciclo es m.

4. $(12\cdots m)^{-1} = (m \cdots 21)$.

5. Si $\alpha = (a_1 \cdots a_{n_1}) \cdots (r_1 \cdots r_{n_s})$ es la descomposición cíclica de α, entonces $\circ(\alpha) = \text{m.c.m.}\{n_1, ..., n_s\}$.

Proposición 1.37 *Si $n \geq 3$, el conjunto de todos los 3-ciclos genera A_n.*

Demostración: Sabemos que todo 3-ciclo es par y que toda permutación par se puede escribir como el producto de un número par de transposiciones. La demostración se termina al observar que $(abc) = (ac)(ab)$ y que $(ab)(cd) = (ab)(bc)(bc)(cd) = (bca)(cdb)$ si a, b, c y d son todos distintos. □

Si en la descomposición de σ aparecen z_i i-ciclos disjuntos, esta vez escribiendo también los 1-ciclos, entonces

$$\circ(C(\sigma)) = \frac{n!}{1^{z_1} z_1! 2^{z_2} z_2! \cdots}, \tag{1.2}$$

como es fácil ver contando permutaciones con la misma descomposición cíclica. De ahí que

$$\circ(Z(\sigma)) = 1^{z_1} z_1! 2^{z_2} z_2! \cdots \tag{1.3}$$

Ejemplos.

1. El número de transposiciones en S_n es
$$\frac{n!}{(n-2)!2} = \frac{n(n-1)}{2}.$$

2. El número de r-ciclos en S_n es
$$\frac{n!}{(n-r)!r}.$$

3. En particular, el número de n-ciclos en S_n es $(n-1)!$ y cada n-ciclo conmuta exactamente con sus potencias.

4. El número de productos de k transposiciones disjuntas en S_n es
$$\frac{n!}{2^k k!(n-2k)!}.$$

5. En S_6 hay $(6 \times 5)/2 = 15$ transposiciones.

6. El número de conjugados de $(12)(34)(56)$ en S_6 es $6!/(2^3 3!) = 15$.

Proposición 1.38 *Si $n \geq 3$, entonces el centro de S_n es trivial. En particular,* $\text{Int}\, S_n \cong S_n$ *para $n \geq 3$.*

Demostración: Si $\sigma \in S_n$ tiene un ciclo de longitud ≥ 3, entonces tenemos que $\sigma = (123\cdots)\cdots \notin Z$ porque σ no conmuta con (12). Aquí y más adelante podemos reemplazar a σ por un conjugado suyo.

Si σ no tiene ciclos de longitud ≥ 3, entonces $\sigma = (12)\cdots \notin Z$ porque no conmuta con (13). □

Proposición 1.39 *Si $n \geq 4$, entonces el centro de A_n es trivial y también* $\operatorname{Int} A_n \cong A_n$.

Demostración: Dado $1 \neq \sigma \in A_n$, ó bien $\sigma = (123\cdots)\cdots$, que no conmuta con (12)(34) ó bien $\sigma = (12)(34)\cdots$, que no conmuta con (123). □

Lema 1.40 *Sea $\sigma \in A_n$, entonces $Z_{S_n}(\sigma) \subseteq A_n \Leftrightarrow$ la descomposición cíclica de σ, con z_i i-ciclos, satisface $z_2 = z_4 = \cdots = 0; z_1, z_3, ... \leq 1$.*

Demostración: $\Rightarrow$: Aquí suponemos que σ conmuta solamente con permutaciones pares; pero σ conmuta con cada uno de sus ciclos, por lo que estos son pares, es decir, de longitud impar. Si σ tiene 2 ciclos de la misma longitud impar i, entonces σ conmuta con el producto impar de i transposiciones que conjuga uno de esos i-ciclos en el otro. Esta es una contradicción.

$\Leftarrow$: Recíprocamente, si la descomposición cíclica de σ es como en el enunciado y $\tau \in Z_{S_n}(\sigma)$, entonces τ necesariamente conmuta con cada ciclo α de σ y esto implica que τ actúa como potencia de α en los elementos que α mueve. Así, τ es par. □

Teorema 1.41 *La clase de conjugación en S_n de un elemento par σ es una clase de conjugación en A_n o bien es la unión de dos clases de conjugación en A_n con igual número de elementos. Esto último sucede exactamente cuando $Z_{S_n}(\sigma) \subseteq A_n$.*

Demostración: Sean $Z(\sigma)$ el centralizador de σ en S_n y $C(\sigma)$ la clase de conjugación de σ en S_n. Entonces

$$\circ(C_{A_n}(\sigma)) = [A_n : Z_{A_n}(\sigma)] = [S_n : Z(\sigma)] = \circ(C(\sigma)),$$

cuando $Z(\sigma) \nsubseteq A_n$. Mientras que

$$\circ(C_{A_n}(\sigma)) = [A_n : Z_{A_n}(\sigma)] = \frac{1}{2}[S_n : Z(\sigma)] = \frac{1}{2} \circ (C(\sigma)),$$

cuando $Z(\sigma) \subseteq A_n$. □

Corolario 1.42 *Si $n \geq 5$, entonces el conjunto de todos los 3-ciclos es una clase de conjugación en A_n.*

Observación. Si $n \geq 5$, entonces todo subgrupo normal $N \lhd A_n$ con $N \neq \{1\}$ que contenga un 3-ciclo será A_n, pues N contendrá toda la clase de conjugación de los 3-ciclos, que generan A_n.

Ejemplos. Calculamos los órdenes de las clases de conjugación de S_5 y de A_5 en la siguiente tabla:

$\mathbf{S_5}$			$\mathbf{A_5}$	
Partición	Paridad	Número de Elementos	Partición	Número de Elementos
5	par	$\frac{5!}{5} = 4! = 24$	5^+	12
4+1	impar	$\frac{5!}{4} = 30$	5^-	12
3+2	impar	$\frac{5!}{2\cdot 3} = 20$	3+1+1	20
3+1+1	par	$\frac{5!}{2!\cdot 3} = 20$	2+2+1	15
2+2+1	par	$\frac{5!}{2!2^2} = 15$	1+1+1+1+1	1
2+1+1+1	impar	$\frac{5!}{3!\cdot 2} = 10$		
1+1+1+1+1	par	1		

De aquí se obtienen las siguientes conclusiones:

1. $p(5) = 7$, el número de particiones de 5 es 7.

2. Si N es un subgrupo normal no trivial de S_5, entonces N es una unión de clases de conjugación de S_5; por tanto $\circ(N) = 1 + 10x + 20y + 15z + 24w + 30t$ con $x, z, w, t \in \{0, 1\}, y \in \{0, 1, 2\}$. Como $N \neq S_5$, se tiene que $x = 0$. Como $\circ(N) \mid 120$, es fácil ver que la única posibilidad es: $\circ(N) = 1 + 15 + 20 + 24 = 60$, que corresponde a $N = A_5$.

3. De manera similar se puede ver que A_5 es simple.

Ejemplo. A_6 es simple.

Demostración: Supogamos que existe $\{1\} \neq N \lhd A_6$. Entonces existe $1 \neq \sigma \in N$. Si σ tiene un punto fijo i, entonces $\sigma \in N \cap H$, donde $H = \{\alpha \in A_6 \mid \alpha(i) = i\} \cong A_5$. Aquí, $(N \cap H) \lhd H$ con H simple, por lo que $(N \cap H) = H$ y entonces $N \supseteq H$ y N contiene un 3-ciclo. Ya que los 3-ciclos generan A_6, se sigue que $N = A_6$.

Si σ no tiene puntos fijos, entonces podemos escribir $\sigma = (12)(3456)$ ó bien $\sigma = (123)(456)$, pues $(12)(34)(56)$ y (123456) son impares. En todo caso, el número de conjugados de σ en A_6 es $\frac{6!}{2\cdot 4} = 90$ ó bien es $\frac{6!}{2!\cdot 3^2} = 40$.

Se concluye que si N no contiene elementos con puntos fijos, entonces $\circ(N) = 1+90a+40b$, con $a, b \in \{0, 1\}$ y $a+b \neq 0$. Como además, $\circ(N)|360$, se obtiene una contradicción. □

Teorema 1.43 *Si $n \geq 5$, entonces A_n es simple.*

Demostración: Sean $\{1\} \neq N \lhd A_n$ y $1 \neq \alpha \in N$. Como A_n no tiene centro, existe un 3-ciclo β que no conmuta con α. Entonces el elemento $\gamma = (\alpha\beta\alpha^{-1})\beta^{-1} \neq 1$, que está en N, es un producto de dos 3-ciclos. Se concluye que N intersecta de manera no trivial a un subgrupo H de A_n isomorfo con A_6 (pues γ mueve cuando más seis puntos), entonces $(N \cap H) \lhd H$, por tanto $(N \cap H) = H$ y N contiene un 3-ciclo. Finalmente, $N = A_n$. □

Corolario 1.44 *Si $n \geq 5$, entonces S_n tiene un único subgrupo normal propio que es A_n.*

Demostración: Si $\{1\} \neq N \lhd S_n$, entonces $(N \cap A_n) \lhd A_n$. Por tanto, $N = S_n$ ó bien $N = A_n$ ó bien $\circ(N) = 2$. Pero no existe $N \lhd S_n$ con 2 elementos, porque estos serían centrales. □

Teorema 1.45 *Si $n \neq 6$, entonces* $\mathrm{Aut}(S_n) = \mathrm{Int}(S_n)$. *Si además $n \geq 3$, entonces* $\mathrm{Aut}(S_n) \cong S_n$.

Demostración: En vista de la Proposición 1.38, solamente se requiere demostrar la primera afirmación.

Si $\alpha \in \mathrm{Aut}(S_n)$, entonces α envía clases de conjugación a clases de conjugación y preserva el orden de los elementos. Por tanto, α envía el conjunto de las transposiciones al conjunto de los productos de k transposiciones disjuntas para algún k.

Como el número de transposiciones es $n(n-1)/2$ mientras que el número de productos de k transposiciones disjuntas es $n!/[2^k k!(n-2k)!]$, estos números son distintos, excepto cuando $k = 1$ ó bien $n = 6$ y $k = 3$.

Ya que estamos suponiendo $n \neq 6$, se tiene que $k = 1$, por lo que α envía transposiciones a transposiciones.

Escribamos $\alpha(1r) = (a_r b_r)$. Si $r \neq 2$, entonces $(1r)(12) = (12r)$ tiene orden 3, por lo que $\alpha(12r) = (a_r b_r)(a_2 b_2)$ también tiene orden 3. Esto implica que $(a_r b_r)$ y $(a_2 b_2)$ tienen un número en común, es decir, que $a_2 = a_r$ ó bien $b_2 = b_r$. Este razonamiento puede repetirse reemplazando r por $s \neq 1, 2, r$.

Queremos ver que a medida que r recorre al conjunto $\{2, 3, ..., n\}$, todas las transposiciones $(a_r b_r)$ tienen un número en común.

Consideremos la posibilidad de que $a_2 = a_r$ y $b_2 = b_s$. Entonces $\alpha(12r) = (a_2 b_r)(a_2 b_2) = (a_2 b_2 b_r)$, mientras que $\alpha(12s) = (a_s b_2)(a_2 b_2) = (a_2 a_s b_2)$, por lo que $\alpha(12s)\alpha(12r) = (a_2 a_s b_2)(a_2 b_2 b_r) = (b_2 b_r a_s)$, un elemento de orden 3, hecho que contradice al orden 2 de $(12s)(12r) = (1s)(2r)$.

Como las transposiciones $(a_r b_r)$ tienen un número en común, podemos escribir $\alpha(1r) = (a_2 b_r)$ para toda $r \in \{2, 3, ..., n\}$. Esto demuestra que $\alpha = i_\sigma$ con σ dada por $\sigma(1) = a_2$ y $\sigma(r) = b_r$ para $r \neq 1$. □

Teorema 1.46 (Wilson) *Si p es un número primo, entonces*

$$(p-1)! \equiv -1 \pmod{p}.$$

Demostración: En el grupo de permutaciones en p símbolos S_p, los elementos de orden p son los p-ciclos. El número de p-ciclos es $(p-1)!$

Por el Teorema de Cauchy-Frobenius sabemos que p divide al número de elementos α tales que $\alpha^p = 1$, que son los elementos de orden p junto con la identidad. Por tanto, $p \mid [(p-1)! + 1]$, que es la conclusión. □

Ejercicios

1. Construya un morfismo inyectivo $f : S_n \to A_{n+2}$.

2. Demuestre que $S_n = \langle (12), (12...n) \rangle = \langle (12), (23), ..., (n-1, n) \rangle$.

3. Demuestre que el grupo A_4 no tiene subgrupos de orden 6.

4. Sea $H < S_n$ tal que H contiene una transposición y un $(n-1)$-ciclo. Demuestre que si H es transitivo, entonces $H = S_n$.

5. Sean G un grupo finito y X un conjunto (finito) en el que G actúa transitivamente. La acción de G en X induce otra acción en la clase de los subconjuntos de X. Decimos que la acción de G en X es **primitiva** cuando no existen más particiones $\{A_i\}$ de X que las triviales (un sólo subconjunto o subconjuntos todos de orden 1) tales que para todo $g \in G$ se tenga $gA_i = A_i$ ó bien $gA_i = A_j$.

   a) Demuestre que si se tiene una acción imprimitiva y $x \in A_i$ con A_i elemento de una partición de imprimitividad, entonces el conjunto $H = \{g \in G \mid gA_i = A_i\}$ es un subgrupo propio de G tal que $G_x \subset H$.

   b) Recíprocamente, demuestre que si existe un subgrupo $H < G$ tal que $G_x \subset H \subset G$, para algún $x \in X$, entonces la acción es imprimitiva con $A = \{hx \mid h \in H\}$ elemento de una partición de imprimitividad.

   c) Demuestre que en la situación de *a*) ó *b*), se tiene que el número de subconjuntos en una partición de imprimitividad es $[G : H]$; y que un elemento A de una partición de imprimitividad satisface $\circ(A) = [H : G_x]$.

1.8 Productos Directos y Semidirectos

Si $G_1, ..., G_n$ son grupos, definimos el **producto directo** $G_1 \times \cdots \times G_n$ como el conjunto $\{(g_1, ..., g_n) \mid g_i \in G_i\}$ con multiplicación

$$(g_1, ..., g_n)(g'_1, ..., g'_n) = (g_1 g'_1, ..., g_n g'_n).$$

Se verifica inmediatamente que esto es un grupo.

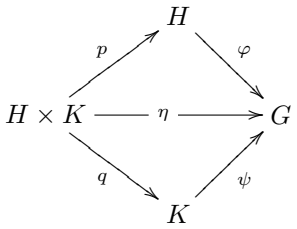
Observaciones. Las siguientes afirmaciones son claras.

1. $G_1 \times G_2 \cong G_2 \times G_1$.
2. $(G_1 \times G_2) \times G_3 \cong G_1 \times (G_2 \times G_3) \cong G_1 \times G_2 \times G_3$ y sus generalizaciones.
3. En $H \times K$, se tiene que $(h, 1)(1, k) = (h, k) = (1, k)(h, 1)$ para todas $h \in H, k \in K$.
4. $H \times \{1\}$ y $\{1\} \times K$ son subgrupos normales de $H \times K$ que generan al producto directo.
5. Si $a \in G_i$ e identificamos a G_i como subgrupo de $G = G_1 \times \cdots \times G_n$, entonces $Z_G(a) = \{(g_1, ..., g_n) \in G \mid g_i \in Z_{G_i}(a)\}$. En particular, $Z(G) = Z_{G_1} \times \cdots \times Z_{G_n}$.

Teorema 1.47 *Si G es un grupo abeliano, $\varphi : H \to G$ y $\psi : K \to G$ son morfismos, entonces existe un morfismo único $\eta : H \times K \to G$ tal que $\eta(h, 1) = \varphi(h)$ y $\eta(1, k) = \psi(k)$, para todas $h \in H, k \in K$.*

Demostración: Es fácil verificar que $\eta : H \times K \to G$ tal que $\eta(h, k) = \varphi(h)\psi(k)$ es un morfismo, necesariamente único porque H y K generan a $H \times K$. □

En la situación del teorema anterior, como se tienen las dos proyecciones $p : H \times K \to H$ y $q : H \times K \to K$, el resultado se expresa como la existencia y unicidad de η, dado el resto del siguiente diagrama conmutativo:



Teorema 1.48 *Sean $H, K \lhd G$ tales que $HK = G$ y $H \cap K = \{1\}$. Entonces $G \cong H \times K$.*

Demostración: Definimos $\varphi : H \times K \to G$ así: $\varphi(h,k) = hk$. Entonces φ es un morfismo porque los elementos de H conmutan con los de K por la Proposición 1.20. Además, φ es suprayectivo por hipótesis.

Tenemos que $\ker\varphi = \{(h,k) \mid hk = 1, h \in H, k \in K\}$; y por lo tanto $(h,k) \in \ker\varphi \Rightarrow h = k^{-1} \in H \cap K = \{1\}$. Así, $\ker\varphi = \{1\}$ y φ es un isomorfismo. □

Teorema 1.49 *Si $G = G_1 \times G_2, H \lhd G_1$ y $K \lhd G_2$, entonces $H \times K \lhd G$ y $G/(H \times K) \cong (G_1/H) \times (G_2/K)$.*

Demostración: Sean $\varphi : G_1 \to G_1/H$ y $\psi : G_2 \to G_2/K$ los morfismos naturales. Definimos $\eta : G \to (G_1/H) \times (G_2/K)$ así: $\eta(a,b) = (\varphi(a), \psi(b))$. Se ve que η es un morfismo suprayectivo con núcleo $H \times K$. □

Observamos en particular que $(G_1 \times G_2)/G_1 \cong G_2$.

Sean G y N dos grupos tales que G actúa en N. Esto quiere decir que tenemos un morfismo $\psi : G \to \operatorname{Aut} N$. Escribimos $g \cdot n$ en lugar de $\psi(g)(n)$ para $g \in G, n \in N$.

Construimos un grupo $H = N \rtimes G$, el **producto semidirecto** de N y G ante la acción dada así: H es como conjunto el producto cartesiano de N y G. La multiplicación en H es como sigue:

$$(x_1, y_1)(x_2, y_2) = (x_1(y_1 \cdot x_2), y_1y_2), \text{ para } x_i \in N, y_i \in G.$$

Aquí $(1,1)$ es la identidad. Verifiquemos la asociatividad:

$$(x_1(y_1 \cdot x_2), y_1y_2)(x_3, y_3) = (x_1(y_1 \cdot x_2)(y_1y_2 \cdot x_3), y_1y_2y_3),$$

mientras que

$$(x_1, y_1)(x_2(y_2 \cdot x_3), y_2y_3) = (x_1\{y_1 \cdot [x_2(y_2 \cdot x_3)]\}, y_1y_2y_3);$$

pero estas expresiones son iguales porque ψ es un morfismo.

Aquí, $(x,y)^{-1} = (y^{-1} \cdot x^{-1}, y^{-1})$, pues por un lado $(x,y)(y^{-1} \cdot x^{-1}, y^{-1}) = (xy \cdot [y^{-1}x^{-1}], yy^{-1}) = (xx^{-1}, yy^{-1}) = (1,1)$; mientras que por el otro, $(y^{-1} \cdot x^{-1}, y^{-1})(x,y) = ([y^{-1} \cdot x^{-1}][y^{-1} \cdot x], y^{-1}y) = (y^{-1} \cdot [x^{-1}x], y^{-1}y) = (1,1)$.

Ejemplo. Sean $N = Z_n = \langle a \rangle$ cíclico de orden n y $G = Z_2 = \langle b \rangle$ cíclico de orden 2. Definimos una acción de G en N así: $b \cdot a^i = a^{-i}$ para toda $i \in \mathbb{N}$. El producto semidirecto $H = N \rtimes G$ es de orden $2n$, se llama **grupo diédrico** y se escribe D_n.

Ejemplo. Sea k un campo. Escribimos k_+ para referirnos al grupo aditivo de k y $k^\star$ para referirnos al grupo multiplicativo $k \setminus \{0\}$. Tenemos que $k^\star$ actúa en k_+ por multiplicación izquierda. El producto semidirecto $k_+ \rtimes k^\star$ es el **grupo afín** $\mathbb{A}_2$. Este grupo también puede ser descrito como el grupo de las **transformaciones afines** $T : k \to k$ de la forma $T(x) = ax + b$ con $a \neq 0$ ante la composición de funciones.

Ejemplo. Cuando la acción de G en N es trivial, es decir, $g \cdot n = n$ para todas $g \in G, n \in N$, entonces $N \rtimes G = N \times G$, el producto directo.

Para todo producto semidirecto $N \rtimes G$, siempre se tiene que la proyección $\pi : N \rtimes G \to G$ es un morfismo suprayectivo con núcleo $N \rtimes \{1\} \cong N$, de manera que $N \rtimes G$ está generado por N y $(\{1\} \rtimes G) \cong G$ con $N \cap G = \{1\}$.

Además se tiene $N \lhd (N \rtimes G)$, de manera que la acción que se obtiene de conjugar N dentro del producto semidirecto con elementos de G coincide con la acción que dio origen al mismo producto:

$$(1, g)(n, 1)(1, g)^{-1} = (g \cdot n, g)(1, g^{-1}) = ((g \cdot n)(g \cdot 1), 1) = (g \cdot n, 1).$$

Recíprocamente, si G es un grupo tal que $G = AB$ con $A \lhd G, B < G$ y $A \cap B = \{1\}$, entonces $G \cong (A \rtimes B)$, con el producto definido por la acción de conjugación en A con elementos de B: En $A \rtimes B$ tenemos que $(a_1, b_1)(a_2, b_2) = (a_1(b_1 a_2 b_1^{-1}), b_1 b_2)$, por lo que $f : A \rtimes B \to G$ dado por $f(a, b) = ab$ es un isomorfismo.

Ejercicios

1. Sean $G_1, ..., G_n$ grupos y $\sigma \in S_n$. Demuestre que $G_1 \times \cdots \times G_n \cong G_{\sigma(1)} \times \cdots \times G_{\sigma(n)}$.

2. Sean $G = G_1 \times G_2$ y $H \lhd G$ tal que $H \cap G_1 = \{1\} = H \cap G_2$. Demuestre que H es abeliano.

3. Verifique que las dos descripciones dadas del grupo afín $\mathbb{A}_2$ dan origen a grupos isomorfos.

4. Sean C un cuadrado con centro en el origen y con lados paralelos a los ejes de coordenadas, R la rotación de 90 grados en sentido contrario a las manecillas del reloj, H la reflección respecto al eje de las x, V la reflección respecto al eje de las y, D la reflección respecto al eje $y = x$ y D' la reflección respecto al eje $y = -x$; y sea $G = \{1, R, R^2, R^{-1}, H, V, D, D'\}$.

   a) Demuestre que G es un grupo ante la composición de funciones.

   b) Demuestre que cualquier función $f : \{a, b\} \to \mathbb{GL}_2$ que satisfaga $f(a) = R$ y $f(b) \in \{H, V, D, D'\} = B$, admite una extensión única a un isomorfismo $\varphi : D_4 \to G$ tal que $\varphi(ab) \in B$.

5. Demuestre que el grupo de cuaternios H y el grupo diédrico D_4 no son isomorfos.

6. *a*) Demuestre que $\mathrm{Aut}(D_4) \cong D_4$.

   b) Exhiba un automorfismo $\alpha \neq 1$ del grupo D_4 tal que el conjunto $\{x \in D_4 \mid \alpha(x) = x^{-1}\}$ sea de orden $6 = (3/4)(\circ(D_4))$.

1.9 Solubilidad y Nilpotencia

Dado un grupo G, definimos el **grupo derivado** ó **conmutador** de G, escrito DG, D^1G, G' ó bien (G,G) como el subgrupo de G generado por todos los conmutadores $aba^{-1}b^{-1}$ de elementos de G.

Más generalmente, para $A, B < G$, definimos

$$(A,B) = \langle aba^{-1}b^{-1} \mid a \in A, b \in B \rangle.$$

Después definimos inductivamente dos sucesiones de subgrupos de G

$$G = D^0G \supseteq D^1G \supseteq D^2G \supseteq \cdots, \tag{1.4}$$

$$G = L_0G \supseteq L_1G \supseteq L_2G \supseteq \cdots, \tag{1.5}$$

así: $D^0G = L_0G = G$, $D^{i+1}G = (D^iG, D^iG)$ y $L_{i+1}G = (G, L_iG)$ para $i \geq 0$.

Aquí, (1.4) es la **serie derivada** de G, mientras que (1.5) es la **serie central descendente** de G.

Es claro que todos los D^iG y todos los L_iG son subgrupos característicos de G y que $D^iG < L_iG$ para toda i.

Se dice que G es **soluble** cuando existe n tal que $D^nG = \{1\}$ y que G es **nilpotente** cuando existe n tal que $L_nG = \{1\}$.

Observaciones. Las siguientes afirmaciones son claras:

1. G abeliano $\Rightarrow$ G nilpotente $\Rightarrow$ G soluble.
2. Todo grupo $D^nG/D^{n+1}G$ es abeliano.
3. Todo grupo $L_nG/L_{n+1}G$ es central en $G/L_{n+1}G$.

Proposición 1.50 *Todo subgrupo y toda imagen homomorfa de un grupo soluble (nilpotente) es soluble (nilpotente).*

Demostración: Si $H < G$, entonces es claro que $D^nH \subseteq D^nG$ y que $L_nH \subseteq L_nG$ para toda $n \geq 0$; por lo que H es soluble (nilpotente) si G lo es.

La otra afirmación es consecuencia de que si $f : G \to H$ es un morfismo suprayectivo de grupos, entonces $f(D^nG) = D^nH$ y $f(L_nG) = L_nH$ para toda $n \geq 0$. □

Proposición 1.51 *Si $N \lhd G$ tal que N y G/N son solubles, entonces G es soluble.*

Demostración: Como G/N es soluble, $D^kG \subseteq N$ para alguna k; pero N es soluble, por tanto existe j tal que $D^{k+j}G \subseteq D^jN = \{1\}$. □

Teorema 1.52 *a) Si $A_1, ..., A_n$ son grupos solubles, entonces $A_1 \times \cdots \times A_n$ es soluble.*

b) Si $A_1, ..., A_n$ son nilpotentes, entonces $A_1 \times \cdots \times A_n$ es nilpotente.

Demostración: *a*) es consecuencia inmediata del teorema anterior, o bien de las inclusiones $D^i(A_1 \times \cdots \times A_n) \subseteq (D^i A_1) \times \cdots \times (D^i A_n)$, válidas para todo i.

b) se sigue de las inclusiones $L_i(A_1 \times \cdots \times A_n) \subseteq (L_i A_1) \times \cdots \times (L_i A_n)$, válidas para todo i. □

Proposición 1.53 *Si G es un grupo con centro Z tal que G/Z es nilpotente, entonces G es nilpotente.*

Demostración: Como G/Z es nilpotente, existe j tal que $L_j G \subseteq Z$; por tanto $L_{j+1}G = \{1\}$. □

Proposición 1.54 *a) Si G es soluble, entonces G contiene un subgrupo normal abeliano distinto de $\{1\}$.*

b) Si G es nilpotente, entonces $Z \neq \{1\}$.

Demostración: *a*) Si $D^n G \neq \{1\}$, pero $D^{n+1}G = \{1\}$; entonces $D^n G$ es normal y abeliano.

b) Si $L_n G \neq \{1\}$, pero $L_{n+1}G = \{1\}$; entonces $L_n G \subseteq Z$. □

Teorema 1.55 *Si $n \geq 5$, entonces S_n no es soluble.*

Demostración: Por un lado, S_n soluble $\Rightarrow$ A_n soluble $\Rightarrow$ $DA_n \lhd A_n$ con $DA_n \neq A_n$. Por otro lado, A_n simple $\Rightarrow$ $DA_n = \{1\} \Rightarrow A_n$ abeliano, que es absurdo. □

Teorema 1.56 *Si G es un grupo de orden p^n con p primo, entonces G es nilpotente.*

Demostración: Procedemos por inducción en n observando que los casos $n = 0, 1$ son ciertos. Como $Z \neq \{1\}$, se tiene $\circ(G/Z) < \circ(G)$, por lo que G/Z es nilpotente y también G lo es. □

Proposición 1.57 *Si G es un grupo nilpotente y $H \neq G$ es un subgrupo, entonces $N_G(H) \neq H$.*

Demostración: Distinguimos dos casos: $Z \nsubseteq H$ y $Z \subseteq H$.

Si $Z \nsubseteq H$, entonces $ZH < G$ tal que $ZH \neq H$ y $ZH \subseteq N_G(H)$.

Si $Z \subseteq H$, entonces procedemos por inducción en k mínimo tal que $L_k G = \{1\}$. En primer lugar, $L_{k-1}(G/Z) = \{1\}$, también $H/Z \neq G/Z$ y así $N_{G/Z}(H/Z) \neq H/Z$. Por lo tanto, $N_G(H) \neq H$. □

Definimos la **serie central ascendente** de G

$$Z_0(G) \subseteq Z_1(G) \subseteq \cdots, \tag{1.6}$$

inductivamente: $Z_0(G) = \{1\}$ y para $i \geq 0$, $Z_{i+1}(G)$ es la imagen inversa en G del centro de $G/Z_i(G)$. Así, es claro que $Z_1(G) = Z$ es el centro de G y que todo $Z_i(G)$ es un subgrupo característico de G.

Teorema 1.58 *G es nilpotente $\Leftrightarrow$ $Z_k = G$, para alguna k. Más precisamente, para cualquier grupo G, si n es el mínimo entero tal que $L_n = \{1\}$, entonces n es también el mínimo entero tal que $Z_n = G$ y recíprocamente.*

Demostración: Suponiendo que $L_n = \{1\}$, demostraremos por inducción en r que $L_{n-r} \subseteq Z_r$. El caso $r = 0$ es claro. Para el paso inductivo, partimos de $L_{n-i} \subseteq Z_i$ sabiendo que $L_{n-(i+1)}/L_{n-i} \subseteq Z(G/L_{n-i})$; como G/Z_i es imagen homomorfa de G/L_{n-i}, se tiene que $L_{n-(i+1)}Z_i/Z_i \subseteq Z(G/Z_i)$. Esto implica que $L_{n-(i+1)}Z_i \subseteq Z_{i+1}$ y que $L_{n-(i+1)} \subseteq Z_{i+1}$.

En particular, se tiene que $G = L_0 \subseteq Z_n$.

Recíprocamente, suponiendo $Z_s = G$ demostraremos por inducción en r que $L_r \subseteq Z_{s-r}$. Inicialmente, $L_0 = G = Z_s$, por lo que suponemos $L_i \subseteq Z_{s-i}$. Como $L_{i+1} = (G, L_i)$, se tiene que $L_{i+1} \subseteq (G, Z_{s-i})$. Por otro lado, $Z_{s-i}/Z_{s-(i+1)} \subseteq Z(G/Z_{s-(i+1)})$, por tanto $(G, Z_{s-i}) \subseteq Z_{s-(i+1)}$. De esta manera, $L_{i+1} \subseteq Z_{s-(i+1)}$.

En particular, $L_s \subseteq Z_0 = \{1\}$. □

Ejercicios

1. De un contraejemplo para cada una de las implicaciones recíprocas de las siguientes implicaciones válidas para un grupo G:

   G cíclico $\Rightarrow$ G abeliano $\Rightarrow$ G nilpotente $\Rightarrow$ G soluble.

2. Demuestre que el grupo S_4 es soluble, mientras que A_5 no lo es.

3. Sean k un campo arbitrario y $G = GL_n(k)$ el grupo multiplicativo de las matrices invertibles $n \times n$ con coeficientes en k. Definimos al grupo $B = \{[a_{ij}] \in G \mid a_{ij} = 0 \text{ cuando } i > j\}$, así como al grupo $U = \{[a_{ij}] \in B \mid a_{ii} = 1 \text{ para toda } i\}$. Demuestre que:

   a) $U \lhd B < G$.

   b) $U = (B, B)$, suponiendo que $n \geq 2$ y que $\circ(k) \geq 4$.

   c) U es nilpotente.

   d) B es soluble.

   e) $B \cong U \ltimes T$.

1.10 Teoremas de Sylow

Lema 1.59 *Sean $p, n, r \in \mathbb{N}$ con p primo, entonces*

$$\binom{p^r n}{p^r} \equiv n \pmod{pn}.$$

Demostración:

$$\binom{p^r n}{p^r} = \frac{(p^r n)!}{p^r!(p^r n - p^r)!} = \frac{p^r n}{p^r}\frac{(p^r n - 1)!}{(p^r - 1)!(p^r n - p^r)!} = n\binom{p^r n - 1}{p^r - 1};\ \text{pero}$$

$$\binom{p^r n - 1}{p^r - 1} = \frac{p^r n - 1}{p^r - 1}\frac{p^r n - 2}{p^r - 2}\cdots\frac{p^r n - p^r + 1}{1} = \prod_{k=1}^{p^r-1}\frac{p^r n - k}{k}$$

$$= \prod_{k=1}^{p^r-1}\left(\frac{p^r n}{k} - 1\right) = (-1)^{p^r-1} + p\frac{m}{t},\ \text{con } m, t \in \mathbb{Z},\ \text{tales que } p \nmid t.$$

Esto es porque una mayor potencia de p divide al numerador que al denominador en cada fracción del producto. Como además $pm/t \in \mathbb{Z}$, se obtiene que $t \mid m$ y así

$$\binom{p^r n - 1}{p^r - 1} \equiv 1 \pmod{p},\ \text{de donde es claro que } \binom{p^r n}{p^r} \equiv n \pmod{pn}.\ \square$$

Teorema 1.60 (Sylow) *Sea G un grupo de orden $p^r m$ con p primo tal que $p \nmid m$ y $r \geq 1$. Si $1 \leq s \leq r$, entonces el número n de subgrupos de orden p^s satisface $n \equiv 1 \pmod{p}$. En particular, tales subgrupos existen.*

Demostración: Sea $\mathcal{C}$ la colección de todos los subconjuntos de G con p^s elementos. G actúa en $\mathcal{C}$ por translación izquierda: $g \cdot X = \{gx \mid x \in X\}$ para todos $g \in G, X \in \mathcal{C}$.

Escribiendo $q = p^{r-s}m$, tenemos que

$$\circ(\mathcal{C}) = \binom{p^s q}{p^s} \equiv q \pmod{pq}, \tag{1.7}$$

por el lema. Así, $p^{r-s+1} \nmid \circ(\mathcal{C})$, por lo que es claro que existe al menos una órbita $\mathcal{O}$ tal que $p^{r-s+1} \nmid \circ(\mathcal{O})$. Al respecto, hacemos dos afirmaciones:

1. Cada órbita contiene cuando más un subgrupo de G.

2. La órbita $\mathcal{O}$ contiene un subgrupo de G si y sólo si $p^{r-s+1} \nmid \circ(\mathcal{O})$, en cuyo caso $H \in \mathcal{O}, H < G \Rightarrow H$ es su propio estabilizador.

Como cosecuencia de estas afirmaciones se tiene que n es el número de órbitas $\mathcal{O}$ que satisfacen $p^{r-s+1} \nmid \circ(\mathcal{O})$. Sean $\mathcal{O}_1, ..., \mathcal{O}_n$ tales órbitas. Además, si $H_i \in \mathcal{O}_i$ es un subgrupo, entonces $\circ(\mathcal{O}_i) = [G : H_i] = p^{r-s}m$, de manera que $\circ(\mathcal{C}) \equiv np^{r-s}m \equiv p^{r-s}m \pmod{p^{r-s+1}}$, usando (1.7). De esta última congruencia se deduce que $n \equiv 1 \pmod{p}$.

Para terminar, pasamos a demostrar las dos afirmaciones.

1. Si $X, g \cdot X \in \mathcal{O}$ son ambos subgrupos, entonces $1 \in g \cdot X \Rightarrow g^{-1} \in X \Rightarrow g \in X \Rightarrow g \cdot X = X$.

2. Para $X \in \mathcal{O}$, escribimos $G^X = \{g \in G \mid g \cdot X = X\}$, el estabilizador de X, que es un subgrupo de G. Si $X < G$, entonces claramente $G^X = X$; pero como $\circ(G) = \circ(\mathcal{O}) \circ (X)$, se tiene que $p^{r-s+1} \nmid \circ(\mathcal{O})$.

   Recíprocamente, si $\mathcal{O}$ es una órbita tal que $p^{r-s+1} \nmid \circ(\mathcal{O})$ y $X \in \mathcal{O}$ es tal que $1 \in X$, entonces $G^X \subseteq X$ y además $\circ(G) = \circ(\mathcal{O}) \circ (G^X)$. Así tenemos que $\circ(G^X) \leq \circ(X) = p^s$ con $p^s \mid \circ (G^X)$. Concluimos que $G^X = X$ es un subgrupo de G. □

Cuando $\circ(G) = p^n m$ con p primo y $p \nmid m$, un **p-subgrupo de Sylow** es un subgrupo de orden p^n. El teorema anterior garantiza la existencia de tales subgrupos.

Lema 1.61 *Si G es de orden p^n con p primo y G actúa en un conjunto finito X tal que $p \nmid \circ(X)$, entonces X tiene un punto fijo.*

Demostración: X es unión disjunta de órbitas $\mathcal{O}$. Para cada $\mathcal{O}$, ó bien $p \mid \circ (\mathcal{O})$ ó bien $\circ(\mathcal{O}) = 1$. Como $p \nmid \circ(X)$, existe al menos una órbita $\mathcal{O}$ tal que $p \nmid \circ(\mathcal{O})$, es decir, $\circ(\mathcal{O}) = 1$. Este es un punto fijo. □

Teorema 1.62 (Sylow) *Sean G un grupo de orden $p^r m$ con p primo tal que $p \nmid m$, P un p-subgrupo de Sylow y H un subgrupo de G de orden p^s. Entonces existe $x \in G$ tal que $H \subseteq xPx^{-1}$. En particular, los p-subgrupos de Sylow son conjugados.*

Demostración: H actúa en G/P por multiplicación izquierda y además $p \nmid \circ(G/P)$. El lema garantiza que existe un punto fijo xP, es decir, que $HxP = xP$; esto es equivalente con $x^{-1}Hx \subseteq P$ y con $H \subseteq xPx^{-1}$. □

Frecuentemente es útil saber el número n de p-subgrupos de Sylow de un grupo G. Para calcular este número, contamos con dos datos: $n \equiv 1 \pmod{p}$ y $n = [G : N(P)]$, si P es uno de ellos, como consecuencia de que estos subgrupos son conjugados.

Ya que $P \subseteq N(P)$, el segundo dato implica que $n \mid [G : P]$. Esta información es suficiente en muchos casos.

Aplicaciones. A continuación veremos 6 aplicaciones de los Teoremas de Sylow y de los métodos usados en sus demostraciones.

Proposición 1.63 *Si $\circ(G) = p^n$ con p primo, entonces existe una cadena*

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\},$$

tal que $G_i \lhd G_{i-1}$ con G_{i-1}/G_i cíclico de orden p para toda i.

Demostración: Para todo $1 \leq m \leq n$ existe $H < G$ con $\circ(H) = p^m$, por lo que también existe una cadena

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\},$$

tal que $\circ(G_i) = p^{n-i}$ para toda i. Como G nilpotente $\Rightarrow N(G_i) \neq G_i$, vemos que $[G_{i-1} : G_i] = p \Rightarrow G_i \lhd G_{i-1}$ con G_{i-1}/G_i siempre cíclico. □

Proposición 1.64 *Si P es un subgrupo de Sylow de G, $N = N_G(P)$ y $H < G$ tal que $N \subseteq H$, entonces $H = N_G(H)$.*

Demostración: Si $x \in G$ normaliza a H, entonces $xPx^{-1} \subseteq H$, así existe $h \in H$ tal que $hxPx^{-1}h^{-1} = P$. Entonces, $hx \in N \subseteq H$ y $x \in H$. □

Proposición 1.65 *Si $\circ(G) = p^n$ con p primo y $\{1\} \neq N \lhd G$, entonces $N \cap Z \neq \{1\}$.*

Demostración: G actúa en N por conjugación. El número de elementos de cada órbita es 1 ó un múltiplo de p. Como $p \mid \circ(N)$, el número de puntos fijos de esta acción es un múltiplo de p. Los puntos fijos son los elementos de $N \cap Z$. □

Proposición 1.66 *Si G es un grupo finito y $H < G$ es de índice igual al mínimo primo p que divide a $\circ(G)$, entonces $H \lhd G$.*

Demostración: H actúa en G/H por multiplicación izquierda. Esta acción no es transitiva porque H es un punto fijo. Si decimos que la acción está dada a través del morfismo $f : H \to S_p$, entonces es claro que $\circ(\operatorname{Im} f) \mid p!$ y que $\circ(\operatorname{Im} f) \mid \circ(H)$, lo que implica que $\circ(\operatorname{Im} f) = 1$ ó p.

Si $\circ(\operatorname{Im} f) = p$, entonces $\operatorname{Im} f$ es un grupo generado por un p-ciclo, en cuyo caso la acción es transitiva. La conclusión es que $\operatorname{Im} f$ es trivial y que todos los puntos aH son puntos fijos, es decir, que $HaH = aH$ para toda $a \in G$, o lo que es lo mismo, $a^{-1}Ha \subseteq H$ para toda $a \in G$. □

Proposición 1.67 *Sea G un grupo de orden pq con p, q primos tales que $p < q$, $p \nmid (q-1)$. Entonces G es cíclico.*

Demostración: Sabemos que existen $H, K < G$ con $\circ(H) = p, \circ(K) = q$. Sean m el número de conjugados de H y n el de K. Entonces $m \equiv 1 \pmod{p}$ y $m \mid q$. Esto implica que $m = 1$ ó q. La igualdad $m = q$ nos conduce a $p \mid (q-1)$, contrario a la hipótesis, por lo que $m = 1$ y $H \lhd G$.

De manera similar, $n \equiv 1 \pmod{q}$, $n \mid p$ y $p < q$ implican $n = 1$ y $K \lhd G$.

Dado que $H \cap K = \{1\}$, se tiene $HK = G$ y entonces $G \cong H \times K$; pero si a genera a H y b genera a K, el producto ab es de orden pq y genera a $H \times K$. Así G es cíclico. □

Teorema 1.68 $\operatorname{Int} S_6$ *es de índice 2 en* $\operatorname{Aut} S_6$.

Demostración: Si α y β son automorfismos externos de S_6, entonces α y β intercambian la clase de conjugación de (12) con la de (12)(34)(56), por lo que $\alpha\beta$ fija a ambas clases de conjugación. Esto implica que $\alpha\beta \in \operatorname{Int} S_6$ como en la demostración del Teorema 1.45, por lo que $[\operatorname{Aut} S_6 : \operatorname{Int} S_6] \leq 2$.

Para terminar, construiremos un automorfismo externo de S_6.

Con este fin, calculamos el número n de 5-subgrupos de Sylow de S_5: Por un lado, $n \equiv 1 \pmod 5$, por otra parte, $n \mid 24 = (5!/5)$; de donde se sigue que $n = 1$ ó 6. Si $n = 1$, entonces habría un único 5-subgrupo de Sylow, normal y conteniendo a los 24 5-ciclos. Esto es absurdo. Por tanto, $n = 6$.

Ahora bien, S_5 actúa transitivamente en el conjunto, de orden 6, de sus 5-subgrupos de Sylow. Sea $\varphi : S_5 \to S_6$ esta acción. Entonces tenemos que $\ker \varphi = \cap N(P)$, la intersección de los normalizadores de los 5-subgrupos de Sylow, es un subgrupo normal de S_5 de índice ≥ 6. La simplicidad de A_5 implica que $\ker \varphi = \{1\}$.

Sea $K = \operatorname{Im} \varphi$. Entonces $K \cong S_5$ y K es un subgrupo transitivo de S_6.

Sea $H = \{\sigma \in S_6 \mid \sigma(6) = 6\}$. Este es un subgrupo no transitivo de S_6 isomorfo con S_5.

Tenemos pues $H \cong K \cong S_5$. Si logramos exhibir un automorfismo $\psi : S_6 \to S_6$ tal que $\psi(H) = K$, entonces tendremos un automorfismo externo, pues H y K no son conjugados: $K = \tau H \tau^{-1} \Rightarrow \tau(6)$ es un punto fijo de K, que contradice la transitividad de K.

El grupo $G = S_6$ actúa por translación izquierda tanto en G/H como en G/K. Sean $\rho : G \to S_{G/H}$ y $\xi : G \to S_{G/K}$ esas acciones. Observamos que

$$\ker \rho = \{x \in G \mid xyH = yH \text{ para toda } y \in G\} = \bigcap_{y \in G} yHy^{-1} = \{1\},$$

porque este núcleo es un subgrupo normal de S_6 con índice > 2, en vista del Corolario 1.44. De manera análoga, se tiene que $\ker \xi = \{1\}$.

Sea $\chi : S_{G/H} \to S_{G/K}$ el morfismo inducido por una biyección arbitraria $\eta : G/H \to G/K$ tal que $\eta(H) = K$. Entonces existe un único automorfismo ψ que hace conmutativo al siguiente diagrama:

$$\begin{array}{ccc} S_6 & \xrightarrow{\rho} & S_{G/H} \\ \psi \downarrow & & \downarrow \chi \\ S_6 & \xrightarrow{\xi} & S_{G/K} \end{array} \quad , \ (\psi = \xi^{-1} \circ \chi \circ \rho).$$

Se concluye que ψ es externo porque $\psi(H) = K$. □

Ejercicios

1. Sean G un grupo finito, P un subgrupo de Sylow y a, b dos elementos del centro de P tales que exista $x \in G$ con $b = xax^{-1}$. Demuestre que existe $y \in N_G(P)$ tal que $b = yay^{-1}$.

2. Sean G un grupo finito, $H \lhd G$ y P un p-subgrupo de Sylow de H. Demuestre que $G = HN_G(P)$.

3. Sean G un grupo finito, $N \lhd G$ y P un p-subgrupo de Sylow de G. Demuestre que $N \cap P$ es un p-subgrupo de Sylow de N.

4. Sean G un grupo finito, p el mínimo primo que divide a $\circ(G)$ y P un p-subgrupo de Sylow de G. Demuestre que $N_G(P) = Z_G(P)$, en caso de que P sea cíclico.

5. Sean $p > q$ números primos. Demuestre que todo grupo de orden $p^n q$ es soluble.

6. Sean $p < q < r$ números primos y G un grupo de orden pqr.

   a) Demuestre que un q-subgrupo de Sylow o un r-subgrupo de Sylow de G es normal; pero que en todo caso G contiene un subgrupo normal H de orden qr.

   b) Demuestre que un r-subgrupo de Sylow de H es característico y que un r-subgrupo de Sylow de G es normal.

   c) Si $q \nmid (r-1)$, entonces un q-subgrupo de Sylow de G también es normal.

7. Demuestre que un grupo finito G es nilpotente si y sólo si G es el producto directo de sus subgrupos de Sylow. También demuestre que G es nilpotente si y sólo si todo subgrupo de Sylow de G es normal.

8. *a*) Demuestre que el orden de los p-subgrupos de Sylow de S_n es $p^{\nu_p(n!)}$ con
$$\nu_p(n!) = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots,$$
   donde $[x]$ significa "el mayor entero contenido en x".

   b) Demuestre que si se escribe
$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_r p^r$$
   con $0 \leq a_i < p$ para todo i, entonces
$$\nu_p(n!) = \sum_{i=1}^{r} a_i (1 + p + \cdots + p^{i-1}).$$

1.11 Series de Composición

Una **serie subnormal** de un grupo G es una cadena de subgrupos

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}. \tag{1.8}$$

tal que $G_{i+1} \lhd G_i$ para toda i. Cuando todas las inclusiones son estrictas, la **longitud** de la serie es el número de ellas.

Un **refinamiento** de una serie subnormal es otra serie que contiene a la primera. Se dice que una serie subnormal es una **serie de composición** cuando no admite refinamientos de estrictamente mayor longitud.

Dos series subnormales

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\},$$

$$G = H_0 \supseteq H_1 \supseteq \cdots \supseteq H_n = \{1\},$$

son **equivalentes** cuando existe una biyección φ del conjunto de los **factores** G_i/G_{i+1} de una serie al conjunto de los factores H_j/H_{j+1} de la otra, de manera que si $\varphi(i) = j$, entonces $G_i/G_{i+1} \cong H_j/H_{j+1}$.

Los factores de una serie de composición se llaman **factores de composición** del grupo. Los factores de composición de un grupo dado, son una colección de invariantes del grupo:

Teorema 1.69 (Jordan-Hölder) *Dos series de composición del mismo grupo son equivalentes.*

Este teorema es consecuencia inmediata del siguiente, por lo que es suficiente demostrar este último.

Teorema 1.70 (Schreier) *Dos series subnormales*
a) $G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_r = \{1\}$,
b) $G = H_0 \supseteq H_1 \supseteq \cdots \supseteq H_s = \{1\}$,
de un grupo arbitrario G poseen refinamientos equivalentes.

Demostración: Procedemos por inducción en s, observando que la conclusión es clara si $r = 1$ ó $s = 1$.

Primero demostraremos el caso $s = 2$ por inducción en r:

Aquí, la segunda serie es $G \supseteq H \supseteq \{1\}$. Sean $A = G_1H$ y $B = G_1 \cap H$, de manera que $A, B \lhd G$.

Como las series $G_1 \supseteq G_2 \supseteq \cdots \supseteq G_r = \{1\}$ y $G_1 \supseteq B \supseteq \{1\}$ son de longitudes $r-1$ y 2, la hipótesis inductiva garantiza que estas series tienen refinamientos equivalentes:

$$G_1 \supseteq \cdots \supseteq G_2 \supseteq \cdots \supseteq \{1\} \rightleftarrows G_1 \supseteq \cdots \supseteq B \supseteq \cdots \supseteq \{1\}. \tag{1.9}$$

Como $A/H \cong G_1/B$ y $A/G_1 \cong H/B$, se tiene la equivalencia siguiente:

$$A \supseteq H \supseteq B \supseteq \{1\} \rightleftarrows A \supseteq G_1 \supseteq B \supseteq \{1\}. \tag{1.10}$$

La serie de la derecha de (1.9) da lugar a un refinamiento de la serie de la derecha de (1.10), para el cual hay un refinamiento equivalente de la serie de la izquierda de (1.10):

$$A \supseteq \cdots \supseteq H \supseteq B \supseteq \cdots \supseteq \{1\} \rightleftarrows A \supseteq G_1 \supseteq \cdots \supseteq B \supseteq \cdots \supseteq \{1\}. \quad (1.11)$$

De (1.9) y de (1.11) se obtiene la equivalencia

$$G \supseteq A \supseteq G_1 \supseteq \cdots \supseteq G_2 \supseteq \cdots \supseteq \{1\} \rightleftarrows$$

$$G \supseteq A \supseteq \cdots \supseteq H \supseteq B \supseteq \cdots \supseteq \{1\},$$

que demuestra el caso $s = 2$.

En el caso general de s arbitrario, primero obtenemos un refinamiento de a) equivalente a un refinamiento de $G \supseteq H_1 \supseteq \{1\}$:

$$G \supseteq \cdots \supseteq G_1 \supseteq \cdots \supseteq G_2 \supseteq \cdots \supseteq G_r = \{1\} \rightleftarrows$$

$$G \supseteq \cdots \supseteq H_1 \supseteq \cdots \supseteq \{1\}. \quad (1.12)$$

Por la hipótesis inductiva, la serie $H_1 \supseteq H_2 \supseteq \cdots \supseteq H_s = \{1\}$ y la subserie $H_1 \supseteq \cdots \supseteq \{1\}$ de la serie de la derecha de (1.12) tienen refinamientos equivalentes:

$$H_1 \supseteq \cdots \supseteq H_2 \supseteq \cdots \supseteq \{1\} \rightleftarrows H_1 \supseteq \cdots \supseteq \{1\}. \quad (1.13)$$

En estas condiciones, el lado derecho de (1.13) produce un refinamiento del lado derecho de (1.12), para el cual existe un refinamiento equivalente de su lado izquierdo:

$$G \supseteq \cdots \supseteq G_1 \supseteq \cdots \supseteq G_2 \supseteq \cdots \supseteq G_r = \{1\} \rightleftarrows G \supseteq \cdots \supseteq H_1 \supseteq \cdots \supseteq \{1\}$$

$$\rightleftarrows G \supseteq \cdots \supseteq H_1 \supseteq \cdots \supseteq H_2 \supseteq \cdots \supseteq \{1\},$$

lo cual concluye la demostración. □

Corolario 1.71 *Si G tiene una serie de composición, entonces toda serie subnormal de G se puede refinar a una serie de composición de G. En particular, todo subgrupo normal de G es parte de una serie de composición.*

Teorema 1.72 *Si G es un grupo soluble finito, entonces G admite una serie de composición $G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$ tal que todo factor de composición G_i/G_{i+1} es cíclico de orden primo.*

Demostración: Si existe una serie subnormal como en el enunciado, entonces es claro que se trata de una serie de composición, pues no es posible refinarla.

Recíprocamente, la serie $G = D^0G \supseteq D^1G \supseteq \cdots \supseteq D^rG = \{1\}$ con cada $D^iG/D^{i+1}G$ abeliano provee un punto de partida que nos permite suponer que G es abeliano.

El Teorema de Cauchy garantiza la existencia de un subgrupo $H < G$ de orden primo. Como $H \lhd G$ y $\circ(G/H) < \circ(G)$, la demostración concluye por inducción en $\circ(G)$. □

El Teorema de Jordan-Hölder afirma que los factores de composición de un grupo dado, son invariantes del grupo. Si nos restringimos a la colección de grupos abelianos, este resultado sigue siendo válido y tenemos la simplificación de que todas las series subnormales son normales. Dada esta restricción, podemos pasar a la colección de los espacios vectoriales sobre un campo fijo k, al restringirnos a los grupos abelianos que admiten sobre ellos una acción del campo que los convierte en espacios vectoriales.

Es fácil verificar que en ese caso, se cumplen los teoremas de isomorfismo análogos a los de grupos, que fueron los ingredientes usados en la demostración del Teorema de Schreier. De manera que si k es un campo y V es un espacio vectorial con bases $\{u_1, ..., u_n\}$ y $\{v_1, ..., v_m\}$, entonces

$$V = \langle u_1, ..., u_n\rangle \supseteq \langle u_1, ..., u_{n-1}\rangle \supseteq \cdots \supseteq \langle u_1\rangle \supseteq (0),$$

$$V = \langle v_1, ..., v_m\rangle \supseteq \langle v_1, ..., v_{m-1}\rangle \supseteq \cdots \supseteq \langle v_1\rangle \supseteq (0),$$

son series de composición. El Teorema de Jordan-Hölder afirma que $n = m$.

Ejercicios

1. Exhiba una serie de composición para S_n con n arbitrario.

2. Usando el Teorema de Jordan-Hölder, demuestre que en los enteros se tiene factorización única.

3. *a*) Demuestre que todo grupo finito tiene una serie de composición.
   b) Demuestre que todo grupo abeliano que admite una serie de composición, es finito.

4. Complete los pasos indicados para construir una demostración del Teorema de Jordan-Hölder para grupos finitos, independiente del Teorema de Schreier:

   Dos series de composición dadas, de un grupo finito G, inician así:

   $$G = G_0 \supseteq G_1 \supseteq \cdots \quad \text{y} \quad G = G_0 \supseteq G_1^* \supseteq \cdots,$$

   Procedemos por inducción en $\circ(G)$.

   a) Si $G_1 = G_1^*$, se tiene la conclusión.

   b) Si $G_1 \neq G_1^*$, se tiene que $G_1 G_1^* = G$; y que

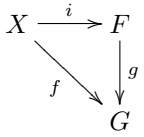
   $$G/G_1 \cong G_1^*/(G_1 \cap G_1^*) \quad \text{y} \quad G/G_1^* \cong G_1/(G_1 \cap G_1^*).$$

   c) Use una serie de composición de $G_1 \cap G_1^*$ para completar a las series $G = G_0 \supseteq G_1 \supseteq G_1 \cap G_1^*$ y $G = G_0 \supseteq G_1^* \supseteq G_1 \cap G_1^*$ hasta tener dos series de composición equivalentes para G.

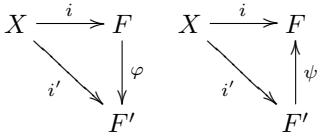
   d) Concluya que las series originalmente dadas son equivalentes.

1.12 Generadores y Relaciones

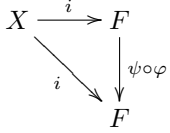
Sean X un conjunto no vacío y F un grupo. Se dice que F es el **grupo libre en X** cuando existe una función inyectiva $i : X \to F$ tal que para toda función $f : X \to G$, donde G sea un grupo, exista un único morfismo de grupos $g : F \to G$ haciendo conmutativo al siguiente diagrama :



Para todo conjunto no vacío X, siempre existe un grupo libre F en X, que es único, pues si F' fuera otro, entonces existirían funciones inyectivas $i : X \to F$, $i' : X \to F'$ y morfismos únicos $\varphi : F \to F'$ y $\psi : F' \to F$ haciendo conmutativos a los diagramas :



Así, $i' = \varphi \circ i,\ i = \psi \circ i'$; y entonces $\psi \circ \varphi$ hace conmutativo al diagrama



El único morfismo con tal propiedad es la identidad en F. Esto implica que $\psi \circ \varphi$ es la identidad en F. De manera similar, $\varphi \circ \psi$ es la identidad en F', por lo que $\varphi : F \to F'$ es un isomorfismo.

La existencia del grupo libre F en X se puede demostrar construyéndolo: Para tal fin se crea un alfabeto con tantas "letras" a como elementos tenga X, junto con nuevos símbolos a^{-1} para cada $a \in X$, más el símbolo 1. El grupo F consistirá de todas las palabras con (un número finito de) letras del alfabeto descrito. La multiplicación en F es yuxtaposición de palabras y se permiten cancelaciones del símbolo 1 y de las parejas aa^{-1} y $a^{-1}a$ donde quiera que ocurran. La palabra vacía es el elemento identidad 1. Falta por verificar la asociatividad de F; y se propone como ejercicio.

Proposición 1.73 *Todo grupo es cociente de un grupo libre.*

Demostración: Si G es un grupo y $A \subseteq G$ es tal que $\langle A \rangle = G$, entonces la inclusión $j : A \to G$ se puede extender de manera única a un morfismo suprayectivo $\eta : F \to G$ con F grupo libre en A. □

En las condiciones de la proposición escribimos $R = \ker \eta$, para tener $G \cong F/R$. Decimos que R es el **grupo de relaciones** de G.

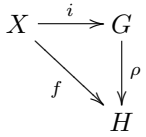
Decimos que A es un **sistema de generadores** de G, mientras que un conjunto B de generadores de R es un **sistema de relaciones** de G. La información anterior es una **presentación** de G.

Ejemplo. El grupo cíclico Z_n admite un generador a y un sistema de relaciones generadas por a^n. Esto se escribe así: $Z_n = \langle a \mid a^n = 1 \rangle$.

Observaciones. Un grupo puede admitir más de una presentación, como es el caso de $Z_{pq} = \langle a \mid a^{pq} = 1 \rangle = \langle a, b \mid a^p = 1, b^q = 1, aba^{-1}b^{-1} = 1 \rangle$, cuando p, q son números primos tales que $p < q, p \nmid (q-1)$, como se vió en la Proposición 1.67.

El grupo libre en un generador es el grupo aditivo $\mathbb{Z}$.

Si F es un grupo libre en X, entonces consideramos al grupo derivado $R = \langle aba^{-1}b^{-1} \mid a, b \in F \rangle$. El cociente $G = F/R$ es el **grupo libre abeliano** en X. Para el grupo G, existe una función inyectiva $i : X \to G$ tal que dada una función $f : X \to H$, donde H sea un grupo abeliano, siempre existe un único morfismo de grupos $\rho : G \to H$ que hace conmutativo al siguiente diagrama :



Esta propiedad la podemos usar para caracterizar al grupo libre abeliano en X o usarla como definición. En todo caso, el grupo libre abeliano en un conjunto X resulta ser isomorfo al producto directo de copias del grupo aditivo $\mathbb{Z}$, requiriéndose tantas copias de $\mathbb{Z}$ como elementos tenga X.

Teorema 1.74 *El grupo simétrico S_n admite la presentación*

$$\langle s_1, ..., s_{n-1} \mid s_i^2 = 1, \text{ para } 1 \leq i \leq n-1, \ (s_i s_{i+1})^3 = 1, \text{ para}$$
$$1 \leq i \leq n-2; \text{ y } (s_i s_j)^2 = 1, \text{ para } 1 \leq i \leq n-3 \text{ con } j > i+1 \rangle.$$

Demostración: Por el ejercicio 1.7.2, $S_n = \langle t_i = (i, i+1), 1 \leq i < n \rangle$; de manera que si L es el grupo libre en los generadores s_i con $1 \leq i < n$, entonces el morfismo $\eta : L \to S_n$ tal que $\eta(s_i) = t_i$ es suprayectivo. El núcleo de η contiene al subgrupo normal R de L generado por $s_i^2 = 1$, para $1 \leq i \leq n-1$, $(s_i s_{i+1})^3 = 1$, para $1 \leq i \leq n-2$; y $(s_i s_j)^2 = 1$, para $1 \leq i \leq n-3$ con $j > i+1$. Sea $G = L/R$. Tenemos que η induce un morfismo suprayectivo $\kappa : G \to S_n$.

Es suficiente ver que κ es inyectivo. El siguiente razonamiento demuestra por inducción en n que $\circ(G) \leq n!$:

Sea $H = \langle s_1, ..., s_{n-2} \rangle < G$ y sea A el conjunto de las siguientes clases laterales:

$$H, Hs_{n-1}, Hs_{n-1}s_{n-2}, ..., Hs_{n-1}s_{n-2} \cdots s_1.$$

Claramente, $\circ(A) \leq n$, por ser un conjunto de clases laterales.

Afirmamos que A es estable ante translación derecha por G. Esto implicará que G está contenido en la unión de las clases laterales en A y por tanto que $\circ(G) \leq n \circ (H)$. Pero en el caso $n = 2$, se ve que $\circ(H) = 1$; por lo que inductivamente se obtendrá $\circ(G) \leq n!$.

Para verificar la afirmación pendiente, veamos la multiplicación derecha por el generador s_j. En primer lugar, esta operación intercambia a la clase lateral $Hs_{n-1} \cdots s_j$ con $Hs_{n-1} \cdots s_{j+1}$.

Si $i > j + 1$, entonces $(Hs_{n-1} \cdots s_i)s_j = Hs_{n-1} \cdots s_i$, pues s_j conmuta con toda s_k entre H y s_j, debido a las relaciones $(s_j s_k)^2 = 1$, válidas para toda $k > j + 1$.

Observemos que $s_j s_{j-1} s_j = s_{j-1} s_j s_{j-1}$, por lo que $i < j$ implica

$$(Hs_{n-1} \cdots s_i)s_j = (Hs_{n-1} \cdots s_j s_{j-1})s_j(s_{j-2} \cdots s_i) =$$

$$(Hs_{n-1} \cdots s_{j-1}s_j)s_{j-1}(s_{j-2} \cdots s_i) = (Hs_{n-1} \cdots s_j)s_{j-1}(s_{j-2} \cdots s_i). \square$$

Ejercicios

1. Demuestre que el grupo diédrico D_n admite la presentación

$$\langle s, t \mid s^n = 1,\ t^2 = 1,\ tst^{-1} = s^{-1} \rangle.$$

2. Demuestre que el grupo de cuaternios H tiene presentación

$$\langle a, b \mid a^4 = 1,\ a^2 = b^2,\ bab^{-1} = a^{-1} \rangle.$$

1.13 Grupos Abelianos Finitamente Generados

Todos los grupos que aparecen en esta sección se suponen abelianos. Adoptamos la notación aditiva, de manera que por ejemplo, el elemento identidad es 0.

Estudiaremos primero la estructura y clasificación de los grupos abelianos finitos para después extender los resultados al caso de los grupos abelianos finitamente generados.

Si A es un grupo abeliano finito y p es un número primo, A_p es el p-subgrupo de Sylow de A, el cual es único y normal, también es característico. El subgrupo A_p consiste de los elementos de A cuyo orden es una potencia de p. Se tiene que $A_p \neq (0) \Leftrightarrow p \mid \circ(A)$.

Teorema 1.75 *Si A es un grupo abeliano finito y $p_1, ..., p_r$ son los primos que dividen al orden de A, entonces $A = A_{p_1} \times \cdots \times A_{p_r}$.*

Demostración: Sabemos que $A_{p_i} \lhd A$ para toda i. Si los subgrupos $A_{p_1}, ..., A_{p_r}$ forman producto directo, entonces $A_{p_1} \times \cdots \times A_{p_r}$ será un subgrupo de A del mismo orden que A, por tanto igual con A.

Así que es suficiente verificar que $A_{p_1} \cap A_{p_2} = (0)$, $(A_{p_1} \times A_{p_2}) \cap A_{p_3} = (0)$, etc. Esto es porque m.c.d.$\{q_1, q_2\} = 1$, m.c.d.$\{q_1 q_2, q_3\} = 1$, etc., donde $q_i = \circ(A_{p_i})$. □

Se dice que un grupo G es **finitamente generado** cuando admite a un conjunto finito como sistema de generadores. Un grupo abeliano finitamente generado es finito si y sólo si todos los elementos de un sistema de generadores son de orden finito.

Observación. Dada una colección de subgrupos cíclicos $A_i = \langle a_i \rangle$ de un grupo abeliano A con $1 \leq i \leq m$, el grupo A es producto directo de sus subgrupos A_i cuando se satisfacen las dos condiciones siguientes:

- Para todo $a \in A$, existen $n_i \in \mathbb{Z}$ tales que $a = n_1 a_1 + \cdots + n_m a_m$.
- $n_1 a_1 + \cdots + n_m a_m = 0$ con $n_i \in \mathbb{Z}$ implica $n_i a_i = 0$ para todo i.

Lema 1.76 *Sea $A = \langle g_1, ..., g_r \rangle$ un grupo abeliano; y sean $c_1, ..., c_r \in \mathbb{N}$ tales que* m.c.d.$\{c_1, ..., c_r\} = 1$. *Entonces existen $h_1, ..., h_r \in A$ tales que $A = \langle h_1, ..., h_r \rangle$ y además $h_1 = c_1 g_1 + \cdots + c_r g_r$.*

Demostración: Procedemos por inducción en $n = c_1 + \cdots + c_r$. Siendo claro el caso $n = 1$, suponemos que $n > 1$ para tener que al menos dos de los números c_i son positivos. Escribimos $c_1 \geq c_2 > 0$, de manera que

$$\text{m.c.d.}\{c_1 - c_2, c_2, ..., c_r\} = 1 \text{ y } (c_1 - c_2) + c_2 + \cdots + c_r < c_1 + c_2 + \cdots + c_r.$$

Dado que $A = \langle g_1, g_1 + g_2, g_3, ..., g_r \rangle$, la hipótesis inductiva garantiza que existen elementos $h_1, ..., h_r \in A$ tales que $A = \langle h_1, ..., h_r \rangle$, donde además $h_1 = (c_1 - c_2) g_1 + c_2 (g_1 + g_2) + \cdots + c_r g_r = c_1 g_1 + \cdots + c_r g_r$. □

Teorema 1.77 *Sea A un grupo abeliano tal que admita un sistema de generadores con r elementos. Entonces A es el producto directo de r grupos cíclicos.*

Demostración: Sea $\{g_1, ..., g_r\}$ un sistema de generadores de A tal que $(\circ(g_1), ..., \circ(g_r))$ sea mínimo en el orden lexicográfico entre todos los sistemas de generadores de A con r elementos. Se afirma que

$$A = \langle g_1 \rangle \times \cdots \times \langle g_r \rangle.$$

Para ver esto, supongamos que existen $a_1, ..., a_r \in \mathbb{Z}$ tales que

$$a_1 g_1 + \cdots + a_r g_r = 0 \text{ sin que } a_i g_i = 0 \text{ para todo } i.$$

También digamos que $0 \leq a_i < \circ(g_i)$ para todo i.

Sea s el mínimo índice i tal que $a_i \neq 0$ y sea $d = \text{m.c.d.}\{a_s, ..., a_r\}$. Escribiendo $a_i = db_i$ para toda i se tiene que $\text{m.c.d.}\{b_s, ..., b_r\} = 1$, por lo que el lema garantiza la existencia de un sistema de generadores de $\langle g_s, g_{s+1}, ..., g_r\rangle$ así: $h_s, h_{s+1}, ..., h_r$ con $h_s = b_s g_s + \cdots + b_r g_r$. De este modo, $dh_s = 0$ y también $A = \langle g_1, .., g_{s-1}, h_s, ..., h_r\rangle$ con la contradicción de que $\circ(h_s) \leq d \leq a_s < \circ(g_s)$. □

Observación. Si en el teorema anterior suponemos que $\circ(A) > 1$ y que el sistema de generadores dado tiene un número mínimo de elementos, entonces obtendremos factores directos no triviales.

El siguiente teorema nos da la estructura y clasificación de los grupos abelianos finitos, ahí usamos la siguiente notación: Dado un entero n, escribimos $\sharp(n)$ para indicar el número de particiones de n, es decir, el número de maneras en que n se puede expresar como suma de enteros positivos. Por ejemplo $\sharp(4) = 5$ porque $4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1$.

Teorema 1.78 *a) Todo grupo abeliano finito A es un producto directo de grupos cíclicos de órdenes potencias de primos.*

b) Los órdenes de los factores directos son únicos.

c) Si $m = p_1^{n_1} \cdots p_r^{n_r}$ con $p_1, ..., p_r$ primos distintos con $n_i > 0$ para $1 \leq i \leq r$, entonces el número de grupos abelianos de orden m no isomorfos entre sí es $\sharp(n_1) \cdots \sharp(n_r)$.

Demostración: El Teorema 1.75 dice que A es el producto directo de sus subgrupos de Sylow, aplicamos el Teorema 1.77 a cada uno de estos subgrupos para obtener *a*).

Como *c*) es consecuencia inmediata de *b*), es suficiente ver que los órdenes de los factores directos de un grupo abeliano A son únicos, suponiendo que $\circ(A) = p^n$, con p primo.

Sea $\varphi : Z_{p^k} \to Z_{p^k}$ el morfismo dado por $\varphi(a) = pa$ para toda a en el grupo cíclico Z_{p^k}. Es inmediato que $\varphi(Z_{p^k}) \cong Z_{p^{k-1}}$. Más generalmente, si G es un grupo abeliano de orden una potencia de p y $\varphi : G \to G$ es multiplicación por p, entonces φ es un morfismo de grupos tal que $\circ(\varphi(G)) = \circ(G)/p^s$, donde s es el número de factores cíclicos de G.

El orden de $\varphi(A)$ no depende de la descomposición de A. Por ello, vemos que dos descomposiciones cualesquiera de A deben tener el mismo número de factores.

Generalizando este razonamiento para $\varphi^i(A)$, primero vemos que el orden de cada factor de $\varphi^i(A)$ es p veces el orden de un factor correspondiente de $\varphi^{i+1}(A)$, para después darnos cuenta de que podemos deducir los órdenes de los factores cíclicos de A a partir de los distintos números $\circ(\varphi^i(A))$. □

Ejemplo. Consideremos $A = Z_{p^4} \times Z_{p^2}$ con p primo.

grupo	descomposición	orden
A	$Z_{p^4} \times Z_{p^2}$	p^6
$\varphi(A)$	$Z_{p^3} \times Z_p$	p^4
$\varphi^2(A)$	Z_{p^2}	p^2
$\varphi^3(A)$	Z_p	p
$\varphi^4(A)$	$\{1\}$	1

En la tabla anterior, la columna central se obtuvo de abajo hacia arriba a partir de las columnas externas así:

Como $\circ(\varphi^3(A)) = p$, es claro que $\varphi^3(A) = Z_p$.

A partir de que $\circ(\varphi^2(A))/\circ(\varphi^3(A)) = p$ se tiene que $\varphi^2(A) = Z_{p^2}$.

De $\circ(\varphi(A))/\circ(\varphi^2(A)) = p^2$, se deduce que $\varphi(A)$ tiene 2 factores directos, cuyos órdenes deben ser p^3 y p.

Finalmente, $\circ(A)/\circ(\varphi(A)) = p^2$ nos indica que A tiene dos factores directos cíclicos de órdenes p^4 y p^2 respectivamente.

Los órdenes de los factores directos de A en el teorema anterior se llaman **divisores elementales** de A.

El siguiente teorema es una versión alternativa de la estructura y clasificación de los grupos abelianos finitos. Primero observemos que si $A = \langle a \rangle$ y $B = \langle b \rangle$ son grupos cíclicos de órdenes m y n respectivamente, con m.c.d.$\{m, n\} = 1$, entonces $A \times B$ también es cíclico, generado por ab, pues

$$(ab)^j = 1 \Rightarrow a^j = 1 = b^j \Rightarrow m \mid j \text{ y } n \mid j, \text{ de manera que } mn \mid j.$$

Teorema 1.79 *Si A es un grupo abeliano finito, entonces $A \cong A_1 \times \cdots \times A_r$ con cada A_i cíclico de orden m_i, donde $m_1 \mid m_2 \mid \cdots \mid m_r$.*

Los números $m_1, ..., m_r$ son únicos.

Demostración: El Teorema 1.78 dice que A es un producto directo de grupos cíclicos de órdenes potencias de primos. Para cada primo p que divida al orden de A, hay al menos un factor directo B_p de A, de orden una máxima potencia de p. El producto directo de los distintos B_p al variar p, es un grupo cíclico B, cuyo orden es el mí nimo común múltiplo de los divisores elementales de A. Además, A es el producto directo de B y de los restantes grupos cíclicos de órdenes potencias de primos, cuyo mínimo común múltiplo divide a $\circ(B)$.

Por inducción en $\circ(A)$, se obtiene una descomposición de A como en el enunciado. Nos falta demostrar la unicidad de los órdenes de los factores de tal descomposición.

Supongamos que $A \cong C_1 \times \cdots \times C_r$, con cada C_i cíclico de orden m_i, donde $m_1 \mid m_2 \mid \cdots \mid m_r$, entonces cada C_i es el producto directo de sus subgrupos de Sylow; y estos son cíclicos. Aplicando la unicidad del teorema anterior, tenemos que los órdenes de estos subgrupos de Sylow son únicos. Ahora bien, m_r es el m.c.d. de estos órdenes, m_{r-1} es el m.c.d. de los órdenes que quedan después de eliminar exactamente una potencia máxima de p para cada primo p que divida a m_r; y así sucesivamente. De esta manera concluimos que los números $m_1, m_2, ..., m_r$ son únicos. □

Los órdenes de los factores directos de A en este último teorema, se llam an **factores invariantes** de A.

Corolario 1.80 *Sea G un grupo abeliano finito tal que toda ecuación de forma $dx = 0$ con $0 < d \in \mathbb{N}$ tenga cuando más d soluciones. Entonces G es cíclico.*

Demostración: Como $G \cong G_1 \times \cdots \times G_r$, con cada G_i cíclico de orden n_i con $n_1 \mid n_2 \mid \cdots \mid n_r$, se ve que todo elemento de G es solución de la ecuación $n_r x = 0$, por lo que $G = G_r$ es cíclico. □

En el siguiente corolario usamos la notación multiplicativa para el grupo abeliano que ahí aparece.

Corolario 1.81 *El grupo multiplicativo de todo campo finito es cíclico.*

Demostración: Sean K un campo finito y $K^\star$ el grupo multiplicativo de los elementos distintos de cero de K. Entonces toda ecuación $x^d = 1$ tiene cuando más d soluciones, ver §2.7; por lo que se satisfacen las hipótesis del corolario anterior; así $K^\star$ es cíclico. □

Regresamos a la notación aditiva para los grupos abelianos.

Suponiendo que A es un grupo abeliano finitamente generado, decimos que el elemento $a \in A$ es de **torsión** cuando a tiene orden finito. Definimos la torsión de A como $\operatorname{tor} A = \{a \in A \mid a \text{ es de torsión}\}$. Es claro que si a y b son de orden finito, entonces $-a$ y $a + b$ también lo son. Esto implica que $\operatorname{tor} A$ es un subgrupo (característico) de A. Se dice que A es **de torsión** cuando $A = \operatorname{tor} A$; y que A es **libre de torsión** cuando $\operatorname{tor} A = \{0\}$.

Observación. Si G es un grupo abeliano, entonces $\varphi : G \to G$ dado por $\varphi(a) = a + a$ para toda $a \in G$, es un morfismo cuya imagen escribimos $2G$.

En la sección anterior vimos que un grupo libre abeliano G es un producto directo de copias de $\mathbb{Z}$. El número de factores de esa descomposición es el **rango** de G. A continuación veremos que este concepto está bien definido. Escribimos $\mathbb{Z}^n$ para denotar el producto directo de n copias de $\mathbb{Z}$.

Teorema 1.82 *Si $\mathbb{Z}^m \cong \mathbb{Z}^n$ con $n, m \in \mathbb{N}$, entonces $m = n$.*

Demostración: Como $\mathbb{Z}^m \cong \mathbb{Z}^n$, se tiene que $\mathbb{Z}^m/2\mathbb{Z}^m \cong \mathbb{Z}^n/2\mathbb{Z}^n$. Aplicando el Teorema 1.49 , obtenemos $\circ(\mathbb{Z}^m/2\mathbb{Z}^m) = 2^m$ y $\circ(\mathbb{Z}^n/2\mathbb{Z}^n) = 2^n$; y de ahí, $2^m = 2^n$, por lo que $m = n$. □

Teorema 1.83 *Sea G un grupo abeliano finitamente generado. Entonces G es el producto directo de su torsión,* $\operatorname{tor} G$*, y de un grupo abeliano libre, cuyo rango es un invariante de G.*

Demostración: Si $G \cong (\operatorname{tor} G) \times B$, con B libre abeliano, entonces el grupo $G/\operatorname{tor} G \cong B$ es libre abeliano; y su rango es un invariante de $G/\operatorname{tor} G$, y por tanto de G. Ahora es suficiente ver que tal descomposición existe.

Aplicando el Teorema 1.77, sabemos que $G \cong \langle a_1\rangle \times \cdots \times \langle a_s\rangle$, donde podemos suponer que $a_1, ..., a_m$ tienen orden infinito, mientras que los elementos $a_{m+1}, ..., a_s$ tienen orden finito.

Escribiendo $B = \langle a_1\rangle \times \cdots \times \langle a_m\rangle$ y $C = \langle a_{m+1}\rangle \times \cdots \times \langle a_s\rangle$, tenemos que $G \cong B \times C$ con B libre de rango m y con C de torsión. Así, es inmediato que $C \subseteq \operatorname{tor} G$. Se afirma que $C = \operatorname{tor} G$.

Supongamos que $a = n_1a_1 + \cdots + n_sa_s \in \operatorname{tor} G$. Entonces existe $0 < n \in \mathbb{N}$ tal que $na = nn_1a_1 + \cdots + nn_sa_s = 0$; pero entonces $nn_1 = \cdots = nn_m = 0$, por lo que $n_1 = \cdots = n_m = 0$, de manera que $a \in C$; demostrando la igualdad afirmada. □

El siguiente corolario es inmediato.

Corolario 1.84 *Sea G un grupo abeliano finitamente generado. Entonces G es libre si y sólo si G es libre de torsión.*

Observación. Si G es un grupo abeliano finitamente generado, entonces $\operatorname{tor} G$ también lo es, por lo que $\operatorname{tor} G$ es un grupo finito abeliano, cuya estructura queda descrita por el Teorema 1.78 o bien por el Teorema 1.79. Como consecuencia de esto, tenemos que el Teorema 1.83 completa el estudio de la estructura y clasificación de los grupos abelianos finitamente generados.

Ejercicios

1. Describa $\operatorname{Hom}(Z_{p^n}, Z_{p^m})$, donde p es un primo.
2. Sea G el producto directo de r copias de Z_p con p primo. Demuestre que $\circ(\operatorname{Aut} G) = (p^r - 1)(p^r - p)\cdots(p^r - p^{r-1})$.
3. Sea A un grupo abeliano finito de orden n y sea m un entero positivo tal que $m \mid n$. Demuestre que A contiene un subgrupo de orden m.
4. Investigue si los grupos $Z_8 \times Z_6 \times Z_{10}$ y $Z_4 \times Z_4 \times Z_{30}$ son isomorfos.
5. Sea A un grupo abeliano para el que existe $n \in \mathbb{N}$ tal que $nA = (0)$. Sea $m \in \mathbb{N}$ tal que m.c.d.$\{m, n\} = 1$. Demuestre que para toda $a \in A$, existe $b \in A$ tal que $a = mb$.

6. Sean A y B grupos abelianos tales que $mA = nB = (0)$ con m y n primos relativos. Describa $\mathrm{Hom}(A, B)$.

7. Encuentre los factores invariantes de $Z_m \times Z_n$, si m y n son enteros positivos.

8. Sea G un grupo abeliano con subgrupos A y B de órdenes a y b respectivamente. Demuestre que G contiene un subgrupo de orden m.c.d.$\{a, b\}$.

9. Demuestre que el grupo aditivo de los racionales $\mathbb{Q}$ no se puede expresar como producto directo de dos subgrupos propios.

1.14 Ejercicios Generales

1. Encuentre tres grupos H, K, G tales que $H \lhd K$, $K \lhd G$, sin que H sea normal en G.

2. Sean G un grupo finito y H un subgrupo propio. Demuestre que

$$\bigcup_{a \in G} aHa^{-1} \neq G.$$

3. Sea G un grupo de orden $2n$ con $n > 1$ impar. Demuestre que G no es simple.

4. Sea G un grupo finito tal que $3 \nmid \circ(G)$ y que $(ab)^3 = a^3b^3$ para todos $a, b \in G$. Demuestre que G es abeliano.

5. Sean G un grupo y $a \in G$. Definimos la "translación izquierda" a_L y la "translación derecha" a_R con a como $a_L(x) = ax$ y $a_R(x) = xa$, para toda $x \in G$. De manera que $a_L, a_R : G \to G$.

   También definimos $G_L = \{a_L \mid a \in G\}$ y $G_R = \{a_R \mid a \in G\}$. Estos son subconjuntos de S_G al ser biyecciones todas las a_L, a_R.

   a) Demuestre que $G_L \cdot \mathrm{Aut}\, G$ es un subgrupo de S_G que contiene a G_R. Este grupo llamado **holomorfo** se escribe $\mathrm{Hol}\, G$.

   b) Demuestre que si G es finito, entonces $\circ(\mathrm{Hol}\, G) = \circ(G) \circ (\mathrm{Aut}\, G)$.

   c) Demuestre que $\mathrm{Hol}(Z_3) \cong S_3$ y que $\mathrm{Hol}(Z_4) \cong D_4$.

6. Sea G un grupo finito provisto de un automorfimo f tal que el conjunto $\{x \in G \mid f(x) = x^{-1}\}$ contiene a más de $(3/4) \circ (G)$ elementos. Demuestre que G es abeliano.

7. Demuestre que
$$\sum_{\substack{\text{m.c.d.}\{i,n\}=1 \\ 1\leq i\leq n}} i \;=\; \frac{1}{2}n\varphi(n).$$

8. Determine con demostración todos los grupos G tales que G es isomorfo a todo subgrupo de G distinto de la identidad.

9. Sea G un grupo abeliano con un número finito de subgrupos. Demuestre que G es finito.

10. Sea G un grupo finito de orden n con k clases de conjugación. Demuestre que la probabilidad de que 2 elementos de G escogidos al azar (pero posiblemente iguales) conmuten entre sí es
$$\frac{k+1}{n+1}.$$

11. *a*) Sean H y K subgrupos solubles de un grupo G con $K \lhd G$. Demuestre que HK es soluble.

    b) Observe que en *a*) no es suficiente que $HK < G$ con H y K solubles, viendo el caso en que $G = A_5, H = A_4$ y $K \cong Z_5$.

    c) Demuestre que si H y K son subgupos normales de G con G/H y G/K solubles, entonces el grupo $G/(H \cap K)$ es soluble.

12. Sea p un número primo y sea
$$H = \{\frac{a}{b} \in \mathbb{Q} \mid \text{ m.c.d.}\{a,b\} = 1,\ b \text{ es una potencia de } p\}.$$
    Considere el morfismo natural $\varphi : \mathbb{Q} \to \mathbb{Q}/\mathbb{Z}$; y sea $\mathbb{Z}(p^\infty) = \varphi(H)$.

    Demuestre que $\mathbb{Z}(p^\infty)$ es isomorfo a cada uno de sus cocientes distintos de cero; y que todos sus subgrupos propios son finitos.

13. El **subgrupo de Frattini** $\Phi(G)$ de un grupo G es la intersección de sus subgrupos propios máximos. Se dice que $g \in G$ es un **no generador** cuando para todo subconjunto $A \subseteq G$ tal que $G = \langle A, g\rangle$, se tenga $G = \langle A\rangle$. Demuestre que $\Phi(G)$ es el conjunto de los no generadores.

Capítulo 2
Anillos

2.1 Definiciones y Primeros Resultados

Se dice que R es un **anillo asociativo con** 1 cuando R es un conjunto equipado con dos operaciones: suma y multiplicación, de manera que ante la suma, R es un grupo abeliano; que además satisface las siguientes condiciones:

1. $a, b \in R \Rightarrow ab \in R$.
2. $a(bc) = (ab)c$ para todos $a, b, c \in R$.
3. $a(b + c) = ab + ac, (b + c)a = ba + ca$ para todos $a, b, c \in R$. Estas propiedades se llaman **distributividad** a la izquierda y a la derecha respectivamente.
4. Existe un elemento $1 \in R$ tal que $1 \neq 0$ y $a1 = 1a = a$ para toda $a \in R$.

Por brevedad, diremos **anillo** en lugar de anillo asociativo con 1. En caso de que $ab = ba$ para todos $a, b \in R$, diremos que R es **conmutativo**.

Ejemplos. Como ejemplos de anillos tenemos los siguientes:

1. Los números enteros $\mathbb{Z}$.
2. Los enteros módulo $n : \mathbb{Z}/n\mathbb{Z}$.
3. Los números racionales $\mathbb{Q}$.
4. Los números reales $\mathbb{R}$.
5. Los números complejos $\mathbb{C}$.
6. El anillo de polinomios $R[X_1, ..., X_n]$ en n variables con coeficientes en el anillo R.
7. El conjunto $M_n(R)$ de las matrices $n \times n$ ante la suma y la multiplicación de matrices, con coeficientes en el anillo R.

8. Los **enteros Gaussianos** $\mathbb{Z}[i] = \{a+bi \mid a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$, con $i^2 = -1$.

9. Los **cuaternios reales** $\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$, donde

   - $(a_1 + b_1 i + c_1 j + d_1 k) + (a_2 + b_2 i + c_2 j + d_2 k) = (a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$
   - $i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j$ y la multiplicación es bilineal.

Un **dominio** es un anillo conmutativo donde vale la **ley de cancelación:**

$$ab = ac, a \neq 0 \Rightarrow b = c.$$

Los ejemplos 1,3,4,5 y 8 son dominios.

Un **anillo de división** es un anillo R tal que para toda $0 \neq a \in R$, existe $b \in R$ tal que $ab = ba = 1$. Un **campo** es un anillo de división conmutativo. Se pide demostrar en el problema 1 que los cuaternios reales son un anillo de división. Los ejemplos 3,4 y 5 son campos.

Decimos que un elemento a de un anillo conmutativo R es **divisor de cero** cuando $a \neq 0$ y existe $0 \neq b \in R$ tal que $ab = 0$. Así, un dominio es un anillo conmutativo sin divisores de cero.

Un elemento $a \in R$ es **unidad** cuando existe $b \in R$ tal que $ab = ba = 1$. Las unidades de R forman un grupo multiplicativo escrito $R^\star$.

Observaciones. En todo anillo R se cumplen las siguientes afirmaciones de fácil verificación:

1. $0a = a0 = 0$ para toda $a \in R$.

2. $a(-b) = (-a)b = -ab$ para todas $a, b \in R$.

3. $(-a)(-b) = ab$ para todas $a, b \in R$.

4. $(-1)a = -a$ para toda $a \in R$.

5. $(-1)(-1) = 1$.

Se dice que T es un **subanillo** de R cuando T es un subconjunto de R que forma un anillo ante las mismas operaciones de R, tal que $1 \in T$, donde 1 es la identidad multiplicativa de R.

Teorema 2.1 *Todo dominio finito D es un campo.*

Demostración: Aquí, $0 \neq a \in D \Rightarrow aD \subseteq D$ y $\circ(aD) = \circ(D)$. Por tanto, $aD = D$ y existe $b \in D$ tal que $ab = 1$. □

Teorema 2.2 *Si p es un número primo, entonces $\mathbb{Z}/p\mathbb{Z}$ es un campo.*

Demostración: Por el teorema anterior, es suficiente ver que $\mathbb{Z}/p\mathbb{Z}$ es un dominio; pero esto es consecuencia de que para elementos $a, b \in \mathbb{Z}$, se tiene que $p \nmid a, p \nmid b \Rightarrow p \nmid ab$. □

Para un anillo dado R, definimos el **centro** Z de R así:

$$Z = \{a \in R \mid ab = ba, \ \forall b \in R\}.$$

El centro de un anillo es siempre un subanillo conmutativo. El centro de un anillo de división es un campo.

Teorema 2.3 (Brauer-Cartan-Hua) *Sean D un anillo de división, Z su centro y K un subanillo de división de D tal que $xKx^{-1} \subseteq K$ para todo $0 \neq x \in D$. Entonces o bien $K \subseteq Z$ ó bien $K = D$.*

Demostración: Supongamos que $K \neq D$. Sean $a \in K, x \in D, x \notin K$. Entonces existe $a_1 \in K$ tal que $xa = a_1 x$. Similarmente, existe $a_2 \in K$ tal que $(1 + x)a = a_2(1 + x)$.

Por tanto, $a = a_2 + (a_2 - a_1)x$; lo que implica $x \in K$ a menos que $a_1 = a_2$. Entonces $a_1 = a_2, (1 + x)a = a_1(1 + x) = a_1 + xa$, de donde se obtiene que $a = a_1$.

Acabamos de ver que todo elemento de K conmuta con todo elemento del complemento de K.

Se afirma ahora que $K \subseteq Z$. En efecto, sean $a, k \in K$ arbitrarios; y $h \notin K$. Entonces a conmuta con h y con $h + k$ porque $h + k \notin K$. Esto implica que a conmuta con k. Así, $a \in Z$. □

Ejercicios

1. Para $\alpha = a + bi + cj + dk \in \mathbb{H}$ con $a, b, c, d \in \mathbb{R}$, definimos

$$\overline{\alpha} = a - bi - cj - dk \ \text{ y } \ N(\alpha) = \alpha\overline{\alpha}.$$

   a) Demuestre que $\overline{\alpha\beta} = \overline{\beta}\overline{\alpha}$, para todos $\alpha, \beta \in \mathbb{H}$.

   b) Demuestre que $N(\alpha) = a^2 + b^2 + c^2 + d^2$ y que $N(\alpha\beta) = N(\alpha)N(\beta)$, siempre que $\alpha, \beta \in \mathbb{H}$.

   c) Encuentre el centro de $\mathbb{H}$.

   d) Demuestre que $\mathbb{H}$ es un anillo de división.

2. Sean R un anillo conmutativo y A una matriz $n \times n$ sobre R. Demuestre que $A \in M_n(R)^\star$ si y sólo si $\det A \in R^\star$.

3. Demuestre que en $\mathbb{Z}/n\mathbb{Z}$, todo elemento distinto de cero es o bien una unidad o bien un divisor de cero.

4. Sean k un campo y R el anillo de las matrices 2×2 sobre k. Demuestre que $(AB - BA)^2$ está en el centro de R para todas $A, B \in R$.

2.2 Funciones Aritméticas

Las **funciones aritméticas** son aquellas definidas en los números naturales $\mathbb{N}$ con valores en un campo, como los números reales $\mathbb{R}$.
Ejemplos. Algunos ejemplos importantes son:

1. La función ϵ dada por:
   - $\epsilon(1) = 1$,
   - $\epsilon(n) = 0$, si $n \neq 1$.
2. $1(n) = 1$ para toda n.
3. $id(n) = n$ para toda n.
4. $\sigma(n) =$ suma de los divisores de n.
5. $\sigma_k(n) =$ suma de las k potencias de los divisores de n.
6. $d(n) =$ número de divisores de n.
7. $\varphi(n)$, la función de Euler.
8. $\mu(n)$, la función de Möbius definida como sigue:
   - $\mu(1) = 1$,
   - $\mu(p^2 q) = 0$, si p es primo y $q \in \mathbb{N}$,
   - $\mu(p_1 \cdots p_r) = (-1)^r$; si $p_1, ..., p_r$ son primos distintos.

Observaciones. Es fácil ver que si $n = p_1^{m_1} \cdots p_r^{m_r}$ es una descomposición prima, entonces

$$d(n) = (m_1 + 1) \cdots (m_r + 1),$$

$$\sigma_k(n) = (1 + p_1^k + \cdots + p_1^{km_1}) \cdots (1 + p_r^k + \cdots + p_r^{km_r}) = \prod_{i=1}^{r} \frac{p_i^{k(m_i+1)} - 1}{p_i^k - 1}.$$

Lema 2.4

$$\sum_{d|n} \mu(d) = \epsilon(n).$$

Demostración: Para $n > 1$, sean $p_1, ..., p_r$ los distintos primos que dividen a n. Consideremos la identidad

$$(1 - p_1) \cdots (1 - p_r) = 1 - p_1 - \cdots - p_r + p_1 p_2 + \cdots;$$

y substituyamos 1 en lugar de cada p_i para obtener $\sum_{d|n} \mu(d) = 0$. □

En el conjunto de las funciones aritméticas definimos la **convolución de Dirichlet** así:

$$(f * g)(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right).$$

Es fácil verificar que el conjunto de las funciones aritméticas provisto de las operaciones suma de funciones y convolución de Dirichlet forma un anillo conmutativo con identidad multiplicativa ϵ.

Con esta notación, $1*1 = d$; mientras que el Lema 2.4 afirma que $1*\mu = \epsilon$.

Teorema 2.5 (Fórmula de inversión de Möbius) *Si g está dada por $g(n) = \sum_{d|n} f(d)$, entonces*

$$f(n) = \sum_{d|n} g(d)\mu\left(\frac{n}{d}\right).$$

Demostración: La hipótesis dice que $g = f*1$; pero sabiendo que $1*\mu = \epsilon$ es claro que las funciones 1 y μ son inversas la una de la otra, por lo que $f = g * \mu$, que es la conclusión. □

Proposición 2.6 *$\sum_{d|n} \varphi(d) = n$, es decir, $\varphi * 1 = id$.*

Demostración: El grupo cíclico Z_n tiene un único subgrupo de orden d isomorfo a Z_d para cada $d|n$. Sea $A_d = \{a \in Z_n \mid \circ(a) = d\}$. De manera que A_d es el conjunto de generadores de Z_d; por tanto, $\circ(A_d) = \varphi(d)$. Como $Z_n = \cup_{d|n} A_d$ es una unión disjunta, se tiene que $n = \sum_{d|n} \varphi(d)$. □

Ejercicios

1. Demuestre que en el anillo de las funciones aritméticas A con valores en un campo, $f \in A^\star \Leftrightarrow f(1) \neq 0$.

2. Demuestre que
$$\varphi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right)d,$$
donde φ es la función de Euler y μ es la función de Möbius.

3. Sea n un entero par. Demuestre que
$$\sum_{d|n} \mu(d)\varphi(d) = 0.$$

4. Suponiendo que para todo $0 < n \in \mathbb{N}$, se tiene que
$$f(n) = \sum_{d|n} \mu(d)g\left(\frac{n}{d}\right);$$
demuestre que $g(n) = \sum_{d|n} f(d)$, también para todo $0 < n \in \mathbb{N}$.

2.3 Morfismos e Ideales

Un **morfismo de anillos** es una función $\varphi : R \to S$ tal que $\varphi(1) = 1$, $\varphi(a+b) = \varphi(a) + \varphi(b), \varphi(ab) = \varphi(a)\varphi(b)$, para todas $a, b \in R$.

El **núcleo** de φ, escrito $\ker \varphi$, es $\{x \in R \mid \varphi(x) = 0\}$. Como φ es en particular un morfismo de grupos abelianos aditivos, sabemos que φ es inyectivo si y sólo si $\ker \varphi = 0$. Además, tenemos que si $\varphi(a) = b$, entonces $\varphi^{-1}(b) = a + ker\ \varphi$.

Un **ideal izquierdo** de R es un subgrupo aditivo J tal que $RJ \subseteq J$, donde $RJ = \{ab \mid a \in R, b \in J\}$. Un **ideal derecho** se define análogamente. Un **ideal bilateral** o simplemente **ideal** es simultaneamente un ideal derecho e izquierdo de R. Excluimos entre los posibles ideales izquierdos, derechos o bilaterales al propio R.

Un anillo es **simple** cuando su único ideal (bilateral) es $\{0\}$.

Es fácil ver que toda intersección de ideales (resp. izquierdos, derechos o bilaterales) es un ideal (resp. izq., der., o bilateral). Así, dado un conjunto $A \subseteq R$, si existe un ideal (resp. izq., der., o bilateral) que lo contiene, entonces la intersección de todos los ideales (resp. izq., der., o bilaterales) que contienen a A es un ideal (resp. izq., der., o bilateral); este es el ideal (resp. izq., der., o bilateral) generado por A, escrito (A) ó bien $(a_1, ..., a_n)$ en caso de que $A = \{a_1, ..., a_n\}$.

Un ideal que admite a un solo elemento como generador se llama **principal**.

Teorema 2.7 *En el anillo $\mathbb{Z}$ todo ideal es principal.*

Demostración: Sea I un ideal de $\mathbb{Z}$. Podemos suponer que $I \neq \{0\}$. Entonces I contiene elementos positivos. Sea a el mínimo elemento positivo de I.

Entonces es claro que $(a) \subseteq I$. Afirmamos que esta inclusión es una igualdad. En efecto, si $b \in I$, entonces el algoritmo euclideano garantiza que existen $q, r \in \mathbb{Z}$ tales que $b = aq + r$ con $0 \leq r < a$; pero $r \in I \Rightarrow r = 0$, de manera que $b \in (a)$. □

Observación. Si $\varphi : R \to S$ es un morfismo de anillos, entonces $\ker \varphi$ es un ideal de R.

Dado un anillo R, existe un único morfismo $f : \mathbb{Z} \to R$, pues $f(1) = 1$. La **característica** de R es aquel $n \in \mathbb{N}$ tal que $\ker f = (n)$. De manera que si $n \neq 0$, entonces n es el mínimo entero positivo tal que toda suma $a + \cdots + a$ con n términos vale cero para todo $a \in R$.

Si $\mathfrak{a}$ es un ideal del anillo R y $r \in R$, entonces $r + \mathfrak{a}$ es una **clase lateral** de $\mathfrak{a}$. Estas clases laterales forman una partición de R. El conjunto de ellas se escribe $R/\mathfrak{a}$.

En $R/\mathfrak{a}$ definimos las operaciones de bloque

$$(r + \mathfrak{a}) + (s + \mathfrak{a}) = (r + s) + \mathfrak{a},\ (r + \mathfrak{a})(s + \mathfrak{a}) = rs + \mathfrak{a}.$$

Teorema 2.8 *Si $\mathfrak{a}$ es un ideal de R, entonces $R/\mathfrak{a}$ es un anillo ante las operaciones de bloque y es también una imagen homomorfa de R.*

Demostración: Para saber que $R/\mathfrak{a}$ es un anillo, es suficiente obsevar que las operaciones de bloque están bien definidas y que coinciden con las operaciones $A + B = \{a + b \mid a \in A, b \in B\}, AB = \{ab \mid a \in A, b \in B\}$ para subconjuntos de R. Entonces es claro que $\varphi : R \to R/\mathfrak{a}$ dado por $\varphi(r) = r + \mathfrak{a}$ es un morfismo suprayectivo de anillos. □

El morfismo $\varphi : R \to R/\mathfrak{a}$ dado por $\varphi(r) = r + \mathfrak{a}$ se llama **natural**. Tenemos las siguientes operaciones para ideales $\mathfrak{a}$ y $\mathfrak{b}$ de un anillo:

- $\mathfrak{a} \cap \mathfrak{b}$, su intersección como conjuntos.
- $\mathfrak{a} + \mathfrak{b} = \{a + b \mid a \in \mathfrak{a}, b \in \mathfrak{b}\}$. Este es el ideal generado por $\mathfrak{a} \cup \mathfrak{b}$.
- $\mathfrak{ab} = \{a_1b_1 + \cdots + a_nb_n \mid a_i \in \mathfrak{a}, b_i \in \mathfrak{b}\}$. Este es el ideal generado por $\{ab \mid a \in \mathfrak{a}, b \in \mathfrak{b}\}$.

El siguiente resultado es tan similar a su análogo para grupos que dejamos su demostración al lector como ejercicio.

Teorema 2.9 *Sea $\varphi : R \to S$ un morfismo suprayectivo de anillos con núcleo I. Entonces $S \cong R/I$. Además, existe una biyección del conjunto de los ideales de R que contienen a I al conjunto de los ideales de S, donde al ideal $\mathfrak{b}$ de S le corresponde el ideal $\mathfrak{a} = \varphi^{-1}(\mathfrak{b})$ de R. En estas condiciones, $R/\mathfrak{a} \cong S/\mathfrak{b}$.*

Si $R_1, ..., R_n$ son anillos, podemos construir su **producto directo** así:

$$\prod_{i=1}^{n} R_i = R_1 \times \cdots \times R_n$$

es el producto directo de grupos abelianos con multiplicación definida por componentes:

$$(a_1, ..., a_n)(b_1, ..., b_n) = (a_1b_1, ..., a_nb_n),$$

para $(a_1, ..., a_n),\ (b_1, ..., b_n) \in R_1 \times \cdots \times R_n$.

Observaciones. Las siguientes afirmaciones son claras:

1. La identidad aditiva del producto es $(0, ..., 0)$.
2. La identidad multiplicativa del producto es $(1, ..., 1)$.
3. Si $\mathfrak{a}_1, ..., \mathfrak{a}_n$ son ideales de un anillo R, entonces tenemos el morfismo

$$f : R \to \prod_{i=1}^{n} R/\mathfrak{a}_i,$$

dado por $f(r) = (r + \mathfrak{a}_1, ..., r + \mathfrak{a}_n)$, cuyo núcleo es $\cap_i \mathfrak{a}_i$.

Ejercicios

1. Sea R un anillo simple. Demuestre que la característica de R es o bien cero o bien un número primo.

2. Sean R un anillo simple y Z su centro. Demuestre que Z es un campo.

3. Sea R un anillo con p elementos donde p es un número primo. Demuestre que $R \cong \mathbb{Z}/p\mathbb{Z}$.

4. Sea R un anillo tal que $a^2 = a$ para toda $a \in R$. (Tales anillos se llaman **Booleanos**). Demuestre que $a + a = 0$ para toda $a \in R$; y que R es conmutativo.

5. Sean k un campo y $M_n(k)$ el anillo de las matrices $n \times n$ sobre k. Demuestre que $M_n(k)$ es simple.

6. Sea $f : \mathbb{H} \to M_2(\mathbb{C})$ la función dada por

$$f(a + bi + cj + dk) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}.$$

   Demuestre que f es un morfismo inyectivo de anillos.

7. Sean R un anillo y $e_1, ..., e_n$ elementos del centro de R tales que $e_i e_j = \delta_{ij} e_i$, $e_1 + \cdots + e_n = 1$. Demuestre que R es un producto directo de n anillos.

2.4 Anillos Conmutativos

En esta sección siempre suponemos que tratamos con anillos conmutativos.

Teorema 2.10 *Un anillo conmutativo R es un campo si y sólo si su único ideal es* (0).

Demostración: Si R es un campo y $0 \neq \mathfrak{a}$ es un ideal, entonces existe $0 \neq r \in \mathfrak{a}$ por lo que $1 = rr^{-1} \in \mathfrak{a}$, que es una contradicción.

Recíprocamente, si (0) es el único ideal de R, entonces dado $0 \neq r \in R$, el "ideal" (r) coincide con R, por lo que existe $s \in R$ tal que $rs = 1$. □

Se dice que un ideal $\mathfrak{p}$ es **primo** cuando $ab \in \mathfrak{p}; a, b \in R \Rightarrow a \in \mathfrak{p}$ ó $b \in \mathfrak{p}$. Esto es equivalente a decir que $R/\mathfrak{p}$ es un dominio.

Se dice que un ideal $\mathfrak{m}$ es **máximo** cuando para todo ideal J tal que $\mathfrak{m} \subseteq J$, se tenga $\mathfrak{m} = J$. Esto es equivalente a decir que $R/\mathfrak{m}$ es un campo.

Observación. Es inmediato que todo ideal máximo es primo.

La terminología anterior está de acuerdo con la situación general:

Si un conjunto X está provisto de una relación de orden parcial $\leq$, se dice que $a \in X$ es **máximo** cuando $a \leq b, b \in X \Rightarrow a = b$. Una **cadena** es un subconjunto $Y \subseteq X$ tal que $c, d \in Y \Rightarrow c \leq d$ ó $d \leq c$. Un elemento $k \in X$ es una **cota superior** para un subconjunto $A \subseteq X$ cuando $a \leq k$ para todo $a \in A$.

Axioma 2.11 (Lema de Zorn) *Si X es un conjunto no vacío ordenado parcialmente, tal que toda cadena de X tiene una cota superior en X, entonces existe un elemento máximo en X.*

Teorema 2.12 *Todo anillo conmutativo R contiene al menos un ideal máximo.*

Demostración: Aplicamos el Lema de Zorn al conjunto $\mathcal{C}$ de los ideales de R ordenados ante $\subseteq$. Esto es posible porque si $\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$ es una cadena de ideales, entonces $\cup_i \mathfrak{a}_i$ es un ideal, pues $1 \notin \cup_i \mathfrak{a}_i$, ($1 \in \cup_i \mathfrak{a}_i \Rightarrow 1 \in \mathfrak{a}_i$ para algún i); y es también una cota superior de la cadena. Se concluye que en $\mathcal{C}$ hay al menos un elemento máximo. □

Corolario 2.13 *En todo anillo conmutativo, todo ideal está contenido en al menos un ideal máximo. En particular, todo elemento que no es unidad está contenido en al menos un ideal máximo.*

Demostración: La misma demostración se aplica al conjunto de los ideales que contienen al ideal dado o al elemento no unidad dado. □

Para un anillo R, su **nilradical** $\mathcal{N}$ es la intersección de todos los ideales primos. El **radical de Jacobson** $\mathcal{J}$ es la intersección de todos los ideales máximos. Es inmediato que $\mathcal{N} \subseteq \mathcal{J}$.

Un elemento $a \in R$ es **nilpotente** cuado existe $n \in \mathbb{N}$ tal que $a^n = 0$. Usamos este concepto para formular una caracterización del nilradical de un anillo. También ofrecemos una caracterización del radical de Jacobson.

Teorema 2.14 *a) El nilradical $\mathcal{N}$ de un anillo R es el conjunto de los elementos nilpotentes de R.*

b) $\mathcal{J} = \{a \in R \mid (1 - ab) \in R^\star$ para toda $b \in R\}$.

Demostración: *a)* Si a es nilpotente y $\mathfrak{p}$ es un ideal primo de R, entonces $0 = a^n \in \mathfrak{p}$, para algún entero positivo n. Por tanto, $a \in \mathfrak{p}$. Así, $a \in \mathcal{N}$.

Recíprocamente, sean $a \in R$ no nilpotente y $\mathcal{C}$ la familia de los ideales I de R tales que $a^n \notin I$ para todo $n \in \mathbb{N}$. Como $(0) \in \mathcal{C}$, es claro que $\mathcal{C} \neq \varnothing$.

Por el Lema de Zorn, $\mathcal{C}$ tiene un elemento máximo $\mathfrak{p}$. Para concluir la demostración, es suficiente ver que $\mathfrak{p}$ es un ideal primo:

Sean $r, s \in R$ tales que $r, s \notin \mathfrak{p}$. Entonces, gracias a la maximalidad de $\mathfrak{p}$, existen $m, n \in \mathbb{N}$ tales que $a^m \in (\mathfrak{p}, r)$ y $a^n \in (\mathfrak{p}, s)$. De manera que podemos escribir $a^m = p + rx$ y $a^n = q + sy$ con $p, q \in \mathfrak{p}; x, y \in R$; y entonces $a^{m+n} \in (\mathfrak{p}, rs)$, lo que implica $rs \notin \mathfrak{p}$, por lo que $\mathfrak{p}$ es primo.

b) Si $a \in \mathcal{J}$ mientras que $(1-ab) \notin R^\star$ para algún elemento $b \in R$, entonces $(1-ab) \in \mathfrak{m}$ para algún ideal máximo $\mathfrak{m}$, de donde se obtiene la contradicción $1 \in \mathfrak{m}$.

Recíprocamente, si $a \notin \mathcal{J}$, entonces existe un ideal máximo $\mathfrak{m}$ tal que $a \notin \mathfrak{m}$; pero entonces $(\mathfrak{m}, a) = R$ y existen $b \in R, c \in \mathfrak{m}$ con $1 = c + ab$. Así, $c = 1 - ab \notin R^\star$. □

A continuación reenunciamos y redemostramos el Corolario 1.81.

Teorema 2.15 *El grupo multiplicativo de todo campo finito es cíclico.*

Demostración: Sea G el grupo multiplicativo de un campo finito. Supongamos que $\circ(G) = n$; y observemos que para todo $d \mid n$ se tiene que la ecuación $x^d = 1$ tiene cuando más d soluciones en G, ver §2.7.

Fijemos un divisor d del orden n. Si existe $a \in G$ de orden d, entonces el grupo cíclico $\langle a \rangle$ es el conjunto de soluciones de $x^d = 1$, por lo que todo elemento de G de orden d está en $\langle a \rangle$. Ahora bien, en ese caso, $\langle a \rangle$ tiene exactamente $\varphi(d)$ elementos de orden d.

Hemos demostrado que para cada $d \mid n$, el número de elementos de orden d es cero ó $\varphi(d)$. Así, $\circ(G)$ es igual a la suma de ciertos números de la forma $\varphi(d)$ para los que existen elementos de orden d en G. Pero la Proposición 2.6 dice que $\sum_{d|n} \varphi(d) = n$. Esto garantiza que existen elementos de orden d para todo divisor de n. En particular, hay elementos de orden n, por lo que G es cíclico. □

Se dice que dos ideales $\mathfrak{a}$ y $\mathfrak{b}$ son **primos relativos** cuando $\mathfrak{a} + \mathfrak{b} = R$.

Teorema 2.16 *Si $\mathfrak{a}_1, ..., \mathfrak{a}_n$ es una colección de ideales primos relativos por parejas, entonces*

$$\bigcap_{i=1}^{n} \mathfrak{a}_i = \prod_{i=1}^{n} \mathfrak{a}_i \quad \text{y} \quad \mathfrak{a}_n + \bigcap_{i=1}^{n-1} \mathfrak{a}_i = R.$$

Demostración: La inclusión $\prod_i \mathfrak{a}_i \subseteq \cap_i \mathfrak{a}_i$ es clara.

Veamos que $\cap_i \mathfrak{a}_i \subseteq \prod_i \mathfrak{a}_i$ por inducción en n:

Si $n = 2$, entonces existen $a_1 \in \mathfrak{a}_1$ y $a_2 \in \mathfrak{a}_2$ tales que $a_1 + a_2 = 1$, por lo que $b \in \mathfrak{a}_1 \cap \mathfrak{a}_2 \Rightarrow b = ba_1 + ba_2 \in \mathfrak{a}_1\mathfrak{a}_2$.

Supongamos que $n > 2$ y que

$$\mathfrak{b} = \bigcap_{i=1}^{n-1} \mathfrak{a}_i = \prod_{i=1}^{n-1} \mathfrak{a}_i.$$

Como existen $x_i \in \mathfrak{a}_i, y_i \in \mathfrak{a}_n$ tales que $x_i + y_i = 1$ para $1 \leq i \leq n-1$, se ve que $x_1 \cdots x_{n-1} = (1-y_1) \cdots (1-y_{n-1}) \equiv 1 \pmod{\mathfrak{a}_n}$ y que $x_1 \cdots x_{n-1} \in \mathfrak{b}$; por tanto $\mathfrak{b} + \mathfrak{a}_n = R$ y entonces $\cap_{i=1}^{n} \mathfrak{a}_i = \mathfrak{a}_n \cap \mathfrak{b} = \mathfrak{a}_n\mathfrak{b} = \prod_{i=1}^{n} \mathfrak{a}_i$. □

Teorema 2.17 (Chino del Residuo) *Sean* $\mathfrak{a}_1, ..., \mathfrak{a}_n$ *ideales del anillo* R, $\varphi : R \to \prod_{i=1}^{n}(R/\mathfrak{a}_i)$ *el morfismo dado por* $\varphi(r) = (r + \mathfrak{a}_1, ..., r + \mathfrak{a}_n)$.

El morfismo φ *es suprayectivo* $\Leftrightarrow$ *los ideales* $\mathfrak{a}_1, ..., \mathfrak{a}_n$ *son primos relativos por parejas. En este caso,* φ *induce un isomorfismo* $(R \,/\, \bigcap_{i=1}^{n} \mathfrak{a}_i) \cong \prod_{i=1}^{n}(R/\mathfrak{a}_i)$.

Demostración: Suponiendo φ suprayectiva, existe $b \in R$ tal que $\varphi(b) = (1, 0, ..., 0)$, es decir, tal que $b \in \mathfrak{a}_2 \cap \cdots \cap \mathfrak{a}_n$ con $b - 1 \in \mathfrak{a}_1$, por lo que $\mathfrak{a}_1 + \mathfrak{a}_i = R$ para toda $i > 1$. De manera similar puede demostrarse que $\mathfrak{a}_i + \mathfrak{a}_j = R$ para toda $i \neq j$.

Recíprocamente, φ es suprayectiva si existen $b_1, ..., b_n \in R$ tales que $\varphi(b_i) = (0, ..., 0, 1, 0, ..., 0)$, donde el 1 está en la posición i. Veamos por ejemplo que existe b_n:

Como $\mathfrak{a}_n + \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_{n-1} = R$, existen $u \in \mathfrak{a}_n$ y $b_n \in \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_{n-1}$ con $u + b_n = 1$. Entonces $\varphi(b_n) = (0, ..., 0, 1)$.

Como $\ker \varphi = \cap_i \mathfrak{a}_i$, se tiene el isomorfismo enunciado. □

Teorema 2.18 *Si* $p > 2$ *es un número primo y* $1 \leq n \in \mathbb{N}$, *entonces el grupo* $(\mathbb{Z}/p^n\mathbb{Z})^\star$ *es cíclico.*

Demostración: Si $a \in \mathbb{Z}$, escribimos a' para la clase de $a \pmod p$, mientras que $\overline{a}$ es la clase de $a \pmod{p^n}$. Sea $\psi : (\mathbb{Z}/p^n\mathbb{Z})^\star \to (\mathbb{Z}/p\mathbb{Z})^\star$ el morfismo dado por $\psi(\overline{a}) = a'$. Claramente, ψ es suprayectivo, con núcleo K de orden p^{n-1}.

Para demostrar el teorema, podemos suponer que $n \geq 2$. Afirmamos que K es cíclico y que $K = \langle \overline{1+p} \rangle$. Veremos por inducción en n, que $(1+p)^{p^{n-2}} \not\equiv 1 \pmod{p^n}$. Lo cual es cierto para $n = 2$. Así, suponemos que $n \geq 3$. A partir de $(1+p)^{p^{n-3}} \not\equiv 1 \pmod{p^{n-1}}$, tenemos que existe $s \not\equiv 0 \pmod p$ tal que $(1+p)^{p^{n-3}} = 1 + sp^{n-2}$; y por tanto

$$(1+p)^{p^{n-2}} = 1 + \binom{p}{1} sp^{n-2} + \binom{p}{2} s^2 p^{2(n-2)} + \cdots + s^p p^{p(n-2)} = 1 + sp^{n-1} + tp^n$$

De manera que $(1+p)^{p^{n-2}} \not\equiv 1 \pmod{p^n}$.

Sea ahora $H = \{\overline{a} \mid \overline{a}^{p-1} = 1\}$. Tenemos que $H < (\mathbb{Z}/p^n\mathbb{Z})^\star$ es tal que $H \cap K = \{1\}$. Así, H y K forman producto directo con $\circ(H) \leq \varphi(p^n)/p^{n-1} = p - 1$.

Ahora bien, para todo $\overline{a} \in (\mathbb{Z}/p^n\mathbb{Z})^\star$, se tiene que $\overline{a}^{p^{n-1}} \in H$ con los elementos $\overline{1}^{p^{n-1}}, \overline{2}^{p^{n-1}}, ..., \overline{(p-1)}^{p^{n-1}}$ todos diferentes, pues sus imágenes ante ψ son: $1', 2', ..., (p-1)'$, por el Teorema de Fermat. Esto implica que $\circ(H) = p - 1$ y que $(\mathbb{Z}/p^n\mathbb{Z})^\star = H \times K$.

El grupo H es cíclico porque su imagen homomorfa $\psi(H) = (\mathbb{Z}/p\mathbb{Z})^\star$ es un grupo cíclico de orden $\circ(H)$. Siendo H y K cíclicos de órdenes primos relativos, se deduce que $H \times K = (\mathbb{Z}/p^n\mathbb{Z})^\star$ también es cíclico. □

Teorema 2.19 *a) Los grupos* $(\mathbb{Z}/2\mathbb{Z})^\star$ *y* $(\mathbb{Z}/4\mathbb{Z})^\star$ *son cíclicos.*
b) Si $3 \leq n \in \mathbb{N}$*, entonces* $(\mathbb{Z}/2^n\mathbb{Z})^\star \cong Z_2 \times Z_{2^{n-2}}$.

Demostración: Solamente *b)* requiere de una demostración. Así, suponemos que $n \geq 3$. Si $a \in \mathbb{Z}$, escribimos a' para la clase de $a(\mathrm{mod}\, 4)$; y $\overline{a}$ para la clase de $a(\mathrm{mod}\, 2^n)$. Sea $\psi : (\mathbb{Z}/2^n\mathbb{Z})^\star \to (\mathbb{Z}/4\mathbb{Z})^\star$ el morfismo dado por $\psi(\overline{a}) = a'$. Claramente, ψ es suprayectivo, cuyo núcleo es el conjunto $K = \{\overline{a} \mid a \equiv 1(\mathrm{mod}\, 4)\}$ de orden 2^{n-2}. Resulta que $K = \langle\overline{5}\rangle$, porque

$$5^{2^{n-3}} = (1+4)^{2^{n-3}} \equiv 1 + 4 \cdot 2^{n-3} \not\equiv 1(\mathrm{mod}\, 2^n).$$

La función $\xi : (\mathbb{Z}/2^n\mathbb{Z})^\star \to \{\pm 1\} \times K$, dada por

$$\xi(\overline{a}) = ((-1)^{(a-1)/2}, (-1)^{(a-1)/2}\overline{a})$$

resulta ser un morfismo de grupos, pues $(-1)^{(a-1)/2} = 1 \Leftrightarrow a \equiv 1(\mathrm{mod}\, 4)$, mientras que $(-1)^{(a-1)/2} = -1 \Leftrightarrow a \equiv -1(\mathrm{mod}\, 4)$. El morfismo ξ es claramente inyectivo; y por tanto, biyectivo. □

Ejercicios

1. Demuestre que en todo anillo conmutativo R vale el **Teorema del Binomio** para todos $n \in \mathbb{N}$; $a, b \in R$:

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i}b^i \quad \text{con} \quad \binom{n}{i} = \frac{n!}{i!(n-i)!}.$$

2. Demuestre que en todo anillo conmutativo existen primos mínimos.

3. Sean A un anillo conmutativo y $a \in A$ un elemento nilpotente. Demuestre que $1 + a$ es invertible.

4. Sean $\mathfrak{a}, \mathfrak{b}, \mathfrak{p}$ ideales de un anillo conmutativo con $\mathfrak{p}$ primo y tales que $\mathfrak{ab} \subseteq \mathfrak{p}$. Demuestre que $\mathfrak{a} \subseteq \mathfrak{p}$ ó $\mathfrak{b} \subseteq \mathfrak{p}$.

5. Sea A un anillo conmutativo tal que para todo $a \in A$, exista un entero $n > 1$ tal que $a^n = a$. Demuestre que todo ideal primo de A es máximo.

6. Sean A un anillo conmutativo y $\mathfrak{p}$ un ideal máximo entre los de forma $\mathrm{an}(a) = \{b \in A \mid ba = 0\}$ para $0 \neq a \in A$. Demuestre que $\mathfrak{p}$ es primo.

7. Exprese $\mathbb{R}[X]/(X^3 + X)$ como producto directo de campos.

8. *a)* Demuestre que hay un número infinito de primos en $\mathbb{Z}$.

   b) Sea n un entero positivo. Demuestre que hay n enteros consecutivos tales que cada uno de ellos es divisible por el cuadrado de un entero mayor que uno.

9. Demuestre que $(\mathbb{Z}/n\mathbb{Z})^\star$ es cíclico si y sólo si $n = 2, 4, p^m, 2p^m$; donde p es un primo impar y $m \geq 1$.

2.5 Localización

En esta sección seguimos tratando con anillos conmutativos.

Se dice que un anillo R es **local** cuando tiene un único ideal máximo $\mathfrak{m}$. En esta situación, los elementos de R que no están en $\mathfrak{m}$ son unidades.

Recíprocamente, si R es un anillo con un ideal I tal que $R \setminus I = R^\star$, entonces R es local con ideal máximo I.

Se dice que $S \subset R$ es un **conjunto multiplicativo** cuando se cumplen las condiciones:

- $1 \in S$.
- $0 \notin S$.
- $a, b \in S \Rightarrow ab \in S$.

El propósito principal de considerar un conjunto multiplicativo S es crear un nuevo anillo de fracciones $S^{-1}R$ a partir de R, donde S es el conjunto de denominadores del nuevo anillo.

Dado un conjunto multiplicativo S de un anillo R, definimos al conjunto $S^{-1}R$ como $\{(a, s) \mid a \in R, s \in S\}$ módulo una relación de equivalencia $\sim$ definida así:

$$(a, s) \sim (a', s') \Leftrightarrow \text{ existe } s_1 \in S \text{ tal que } s_1(s'a - sa') = 0. \tag{2.1}$$

La clase de equivalencia de (a, s) se escribe a/s.

El conjunto $S^{-1}R$ adquiere estructura de anillo ante las operaciones

$$\frac{a}{s} + \frac{a'}{s'} = \frac{s'a + sa'}{ss'}, \ \frac{a}{s}\frac{a'}{s'} = \frac{aa'}{ss'}. \tag{2.2}$$

No es difícil ver que estas operaciones están bien definidas y que cumplen las condiciones para formar un anillo. A continuación verificamos que las operaciones están bien definidas:

Si $a'/s' = a''/s''$, es porque existe $s_1 \in S$ tal que $s_1(s''a' - s'a'') = 0$ y entonces $s_1(s''saa' - s'saa'') = 0$, por lo que

$$\frac{aa'}{ss'} = \frac{aa''}{ss''};$$

y la multiplicación está bien definida.

La suma también está bien definida, pues si $a'/s' = a''/s''$, y existe $s_1 \in S$ tal que $s_1(s''a' - s'a'') = 0$, entonces se tiene $0 = s_1(s''ssa' - s'ssa'') = s_1[ss''(sa' + s'a) - ss'(sa'' + s''a)]$, lo que garantiza que

$$\frac{sa' + s'a}{ss'} = \frac{sa'' + s''a}{ss''}.$$

Un ejemplo muy importante de todo lo anterior se da cuando se tienen un anillo conmutativo R y un ideal primo $\mathfrak{p}$, pues entonces $S = R \setminus \mathfrak{p}$ es

un conjunto multiplicativo; en ese caso $S^{-1}R$ se escribe $R_{\mathfrak{p}}$ y se llama el anillo localizado de R en $\mathfrak{p}$.

Si S es un conjunto multiplicativo de R, entonces $\varphi : R \to S^{-1}R$ dado por $\varphi(a) = a/1$ para toda $a \in R$ es un morfismo de anillos, llamado **natural**. Este morfismo satisface $\varphi(S) \subseteq (S^{-1}R)^{\star}$.

Además, $\ker \varphi = \{a \in R \mid \text{existe } s \in S \text{ tal que } sa = 0\}$. De manera que si R es un dominio, entonces φ es inyectivo.

Cuando R es un dominio y $S = R \setminus (0)$, entonces $S^{-1}R$ es un campo, llamado **campo de fracciones** de R. De esta manera, todo dominio puede ser extendido a un campo. Más generalmente, si R es un anillo conmutativo y S es el conjunto de los no divisores de cero de R, entonces $S^{-1}R$ es el **anillo total de fracciones** de R con $\varphi : R \to S^{-1}R$ inyectivo.

Ejemplo. Cuando $R = \mathbb{Z}$ y $S = \mathbb{Z} \setminus (0)$, entonces $S^{-1}R = \mathbb{Q}$ y $\varphi : \mathbb{Z} \to \mathbb{Q}$ es la inclusión usual.

Dados un anillo R, un ideal I y un subconjunto A de R, definimos el **transportador** $(I : A) = \{r \in R \mid rA \subseteq I\}$, esto es un ideal de R que contiene a I.

Si I es un ideal de R, S es un conjunto multiplicativo y $\varphi : R \to S^{-1}R$ es el morfismo natural, entonces $S^{-1}I$ denota al ideal de $S^{-1}R$ generado por $\varphi(I)$.

Proposición 2.20 *Si $f : R \to R'$ es un morfismo de anillos y $\mathfrak{q}$ es un ideal primo de R', entonces $\mathfrak{p} = f^{-1}(\mathfrak{q})$ es un ideal primo de R.*

Demostración: Todo es consecuencia de que el morfismo f induce un morfismo inyectivo $\overline{f} : R/\mathfrak{p} \hookrightarrow R'/\mathfrak{q}$. □

Observación. El resultado anterior es falso para ideales máximos, como puede verse para el caso de (0) en la inclusión $\mathbb{Z} \subseteq \mathbb{Q}$.

Teorema 2.21 *Sean S un conjunto multiplicativo del anillo conmutativo R y $\varphi : R \to S^{-1}R$ el morfismo natural. Entonces:*

a) Todo ideal $\mathfrak{b}$ de $S^{-1}R$ es de la forma $S^{-1}\mathfrak{a}$, donde $\mathfrak{a} = \varphi^{-1}(\mathfrak{b})$ es un ideal de R.

b) Para todo ideal $\mathfrak{a}$ de R, se tiene que

$$\varphi^{-1}(S^{-1}\mathfrak{a}) = \bigcup_{s \in S} (\mathfrak{a} : s).$$

En particular, $S^{-1}\mathfrak{a} = S^{-1}R \Leftrightarrow S \cap \mathfrak{a} \neq \varnothing$.

c) Los ideales primos $S^{-1}\mathfrak{p}$ de $S^{-1}R$ están en correspondencia biunívoca $\mathfrak{p} \leftrightarrow S^{-1}\mathfrak{p}$ con los ideales primos $\mathfrak{p}$ de R que son disjuntos de S.

Demostración: *a)* Sean $\mathfrak{b}$ un ideal de $S^{-1}R$ y $\mathfrak{a} = \varphi^{-1}(\mathfrak{b})$. Entonces es claro que $\mathfrak{a}$ es un ideal de R y que $S^{-1}\mathfrak{a} \subseteq \mathfrak{b}$. Recíprocamente, si $x/s \in \mathfrak{b}$, entonces $x/1 \in \mathfrak{b}$ y por tanto $x \in \mathfrak{a}$. Así, $x/s \in S^{-1}\mathfrak{a}$, por lo que $\mathfrak{b} \subseteq S^{-1}\mathfrak{a}$.

b) La segunda afirmación es consecuencia inmediata de la primera, que es la que demostramos en seguida.

Si $x \in \cup_{s\in S}(\mathfrak{a} : s)$, es porque existen $s \in S, a \in \mathfrak{a}$ tales que $xs = a$, por lo que $(a/s) = (x/1)$, de manera que $x \in \varphi^{-1}(S^{-1}\mathfrak{a})$. Recíprocamente, suponiendo que $x \in \varphi^{-1}(S^{-1}\mathfrak{a})$, se tiene $(x/1) = (a/s)$ con $s \in S, a \in \mathfrak{a}$, por lo que existe $t \in S$ tal que $t(sx - a) = 0$, de manera que $(ts)x \in \mathfrak{a}$; y así $x \in \cup_{s\in S}(\mathfrak{a} : s)$.

c) Ya sabemos, Proposición 2.20, que si $\mathfrak{q}$ es un ideal primo de $S^{-1}R$, entonces $\mathfrak{p} = \varphi^{-1}(\mathfrak{q})$ es un ideal primo de R. Acabamos de ver en el inciso *a*), que $\mathfrak{q} = S^{-1}\mathfrak{p}$, por lo que $\mathfrak{p} \cap S = \varnothing$, debido al inciso *b*).

Recíprocamente, sea $\mathfrak{p}$ un ideal primo de R tal que $\mathfrak{p} \cap S = \varnothing$. Entonces $\varphi^{-1}(S^{-1}\mathfrak{p}) = \mathfrak{p}$, pues $(\mathfrak{p} : s) = \mathfrak{p}$ para todo $s \in S$.

Si $(a/s), (b/t) \in S^{-1}R$ son tales que $(ab/st) \in S^{-1}\mathfrak{p}$, es porque existen elementos $c \in \mathfrak{p},\ s' \in S$ con $(ab/st) = (c/s')$, de manera que existe $t' \in S$ con $t'(s'ab - stc) = 0$, lo que implica que $ab \in \mathfrak{p}$. Finalmente, $(a/s) \in S^{-1}\mathfrak{p}$ ó bien $(b/t) \in S^{-1}\mathfrak{p}$. Así, $S^{-1}\mathfrak{p}$ es primo. □

Como aplicación inmediata de lo anterior, tenemos que si $\mathfrak{p}$ es un ideal primo de un anillo R, entonces $R_{\mathfrak{p}}$ es un anillo local con ideal máximo $\mathfrak{p}R_{\mathfrak{p}}$.

Los conjuntos multiplicativos también pueden ser utilizados para producir ideales primos, como veremos a continuación.

Decimos que un ideal J de un anillo R es máximo respecto a exclusión de un conjunto $C \subseteq R$ cuando J es máximo entre los ideales I de R tales que $I \cap C = \varnothing$.

Proposición 2.22 *Sea S un conjunto multiplicativo del anillo R. Entonces todo ideal $\mathfrak{p}$ máximo con respecto a exclusión de S es primo.*

Demostración: El Lema de Zorn garantiza la existencia de tales ideales.

Si $a, b \in R$ satisfacen $a, b \notin \mathfrak{p}$, entonces la maximalidad de $\mathfrak{p}$ implica que $(\mathfrak{p}, a) \cap S \neq \varnothing$ y que $(\mathfrak{p}, b) \cap S \neq \varnothing$, por lo que existen elementos $p, q \in \mathfrak{p}; s_1, s_2 \in S; x, y \in R$ tales que $s_1 = p + xa$ y $s_2 = q + yb$, de manera que $s_1 s_2 \in (\mathfrak{p}, ab)$, y por tanto $ab \notin \mathfrak{p}$. □

Corolario 2.23 *En todo anillo conmutativo, el conjunto de los divisores de cero es una unión de ideales primos.*

Demostración: Sean C el conjunto de los divisores de cero y $S = R \setminus C$ el conjunto de los no divisores de cero.

Entonces S es un conjunto multiplicativo y $\cup\mathfrak{p} \subseteq C$, al tomar la unión de los ideales máximos respecto a exclusión de S, que son primos.

Si $a \in C$, entonces el ideal (a) contiene solamente a cero y a divisores de cero y está contenido en algún ideal $\mathfrak{p}$ de la unión, por el Lema de Zorn, por lo que $C \subseteq \cup\mathfrak{p}$. □

Ejercicios

1. Describa los subanillos de $\mathbb{Q}$.

2. Sea R un anillo conmutativo con un único ideal primo $\mathfrak{p}$.

   a) Demuestre que todo divisor de cero de R es nilpotente.

   b) Demuestre que la característica de R es cero o bien una potencia de un primo.

3. Demuestre que toda imagen homomorfa de un anillo local es un anillo local.

4. **Lema de Nakayama**. Sean R un anillo local con ideal máximo $\mathfrak{m}$ y sea $\mathfrak{a} = (a_1, ..., a_n)$ un ideal tal que $\mathfrak{m}\mathfrak{a} = \mathfrak{a}$. Entonces $\mathfrak{a} = 0$. Demuestre este resultado completando los siguientes pasos:

   a) Proceda por inducción en el mínimo número n de generadores $a_1, ..., a_n$ de $\mathfrak{a}$.

   b) Escriba $a_n = c_1a_1 + \cdots + c_na_n$, con $c_1, ..., c_n \in \mathfrak{m}$.

   c) Concluya que a_n es redundante.

2.6 Anillos Euclideanos, Principales y de Factorización Única

Se dice que un dominio R es un **anillo euclideano** cuando existe una función $\delta : R \setminus \{0\} \to \mathbb{N}$ que satisface las siguientes condiciones:

1. $a \mid b \Rightarrow \delta(a) \leq \delta(b)$.

2. Dados $a, b \in R \setminus \{0\}$, existen $q, r \in R$ tales que $a = bq + r$ con $r = 0$ ó bien $\delta(r) < \delta(b)$.

Un dominio R se llama **principal** cuando todos sus ideales I son **principales**, es decir, de la forma $I = (a)$ para algún $a \in R$.

Teorema 2.24 *Todo anillo euclideano es principal.*

Demostración: Sean $I \neq (0)$ un ideal y $0 \neq a \in I$ tal que $\delta(a) = \min\{\delta(b) \mid b \in I\}$. Se afirma que $I = (a)$. Para $c \in I$ arbitrario, existen $q, r \in R$ tales que $c = aq + r$, donde $r = 0$ ó bien $\delta(r) < \delta(a)$. Como $r = c - aq \in I$, se tiene que $r = 0$; y así $c \in (a)$. Así, $I = (a)$. □

Observación. Si b no es una unidad en el anillo euclideano D, entonces para toda $a \in D$, se tiene que $\delta(a) < \delta(ab)$, pues $\delta(a) = \delta(ab) \Rightarrow (a) = (ab)$ por un argumento similar al de esta última demostración.

Dos elementos a, b de un anillo conmutativo son **asociados** cuando existe una unidad u tal que $a = ub$. En un dominio, un elemento p es **irreducible** cuando no es unidad; pero $p = ab$ implica que a es unidad ó b es unidad.

En un dominio, d es un **máximo común divisor** de r y s cuando

- $d \mid r,\ d \mid s$.
- $(c \mid r,\ c \mid s) \Rightarrow c \mid d$.

Es inmediato que si d y d' son máximos comunes divisores de r y s, entonces d y d' son asociados, por lo que se tiene la igualdad de ideales $(d) = (d')$. Convenimos que el máximo común divisor de r y s es cualquier generador de este ideal; y esto lo escribimos así: $(r, s) = d$. Los elementos r y s son **primos relativos** cuando $(r, s) = 1$.

Observación. En un anillo principal R, dos elementos dados a y b siempre tienen un máximo común divisor: Existe $d \in R$ tal que vale la igualdad de ideales $(d) = (a, b)$, de manera que claramente $d \mid a,\ d \mid b$; además existen $\lambda, \mu \in R$ tales que $d = \lambda a + \mu b$, por lo que $(c \mid a,\ c \mid b) \Rightarrow c \mid d$.

Lema 2.25 *Sea R principal. Si $p \mid ab$ y $(p, a) = 1$, entonces $p \mid b$.*

Demostración: Como existen $\lambda, \mu \in R$ con $1 = \lambda p + \mu a$, se tiene que $b = \lambda pb + \mu ab$. Así, $p \mid b$. □

Un dominio D es **de factorización única** cuando todo elemento $a \in D$ que no es unidad, se puede escribir como $a = p_1 \cdots p_r$, con todo p_i irreducible; y además si $a = q_1 \cdots q_s$ es otra expresión con todo q_j irreducible, entonces $r = s$ y existen $u_1, ..., u_r$ unidades y una permutación $\sigma \in S_r$ tales que $q_i = u_i p_{\sigma(i)}$ para toda $1 \leq i \leq r$. Aquí, $u_1 \cdots u_r = 1$.

Lema 2.26 *Sea R un dominio principal. Si p es irreducible con $p \mid ab$, entonces $p \mid a$ ó $p \mid b$.*

Demostración: Supongamos que $p \nmid a$. Por el lema anterior, es suficiente ver que $(p, a) = 1$; pero (p, a) es un divisor de p no asociado de p. Así, (p, a) es una unidad. □

Teorema 2.27 *Todo dominio principal R es de factorización única.*

Demostración: Veamos primero que toda no unidad de $R \setminus \{0\}$ admite una factorización como producto de irreducibles.

Sea Σ el conjunto de ideales $I = (a) \neq (0)$ de R tales que a no es un producto de irreducibles. Supongamos que $\Sigma \neq \varnothing$. Si $(a_1) \subseteq (a_2) \subseteq \cdots$ es una cadena de ideales en Σ, entonces $\cup(a_i)$ es un ideal; y existe $b \in R$ con $\cup(a_i) = (b)$. También existe $n \in \mathbb{N}$ tal que $b \in (a_n)$, y por tanto, $\cup(a_i) = (a_n) = (a_{n+1}) = \cdots$. Así, (a_n) es una cota superior de la cadena.

Aplicando el Lema de Zorn, se obtiene un elemento (r) máximo de Σ. En estas condiciones r no es irreducible y existen $s, t \in R$ no unidades con $r = st$. El ideal (r) está contenido propiamente en cada uno de los ideales (s) y (t), que por ello, no están en Σ. Esto implica que s y t son productos de irreducibles y entonces r también lo es. Esta contradicción demuestra la existencia de factorizaciones.

Para ver la unicidad, supongamos que $p_1 \cdots p_r = q_1 \cdots q_s$ con todos p_i, q_j irreducibles. Usando varias veces el lema anterior, se ve que p_1 divide a q_j para alguna j. Digamos que $p_1 \mid q_1$, para obtener $q_1 = u_1 p_1$, $p_2 \cdots p_r = (u_1 q_2) \cdots q_s$, donde u_1 es una unidad. La demostración concluye por inducción (en r ó en s). □

Corolario 2.28 *Todo anillo euclideano es un dominio de factorización única.*

Proposición 2.29 *En un dominio principal R, un ideal (a) es máximo si y sólo si el elemento a es irreducible.*

Demostración: Si a es irreducible, y $(a) \subseteq (b) \neq R$, entonces $a = bc$ con $c \in R$. Por tanto, c es una unidad y $(a) = (b)$.

Recíprocamente, si (a) es un ideal máximo, entonces (a) es primo. Si $a = bc$, queremos ver que b ó c es unidad. Como (a) es primo, $b \in (a)$ ó bien $c \in (a)$. En el primer caso, $b = as$ y también $a = a(sc)$, por lo que c es unidad. El otro caso es análogo. □

Ejemplos.

1. $\mathbb{Z}$ es un anillo euclideano con $\delta(n) = |n|$, el valor absoluto de n.

2. El anillo de polinomios $k[X]$ en una variable, con coeficientes en un campo k, es un anillo euclideano con $\delta(f(X)) =$ grado de f.

3. El anillo de polinomios $\mathbb{Q}[X, Y]$ es un dominio de factorización única, como veremos en la siguiente sección; pero no es principal: El ideal (X, Y) no es principal, pues si existiera $f \in \mathbb{Q}[X, Y]$ con $(X, Y) = (f)$, entonces tendríamos $f|X$, $f|Y$ y así $f = \pm 1$, que implica la absurda igualdad $(X, Y) = \mathbb{Q}[X, Y]$.

4. Sea $R = \mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$. Claramente, R es un dominio. Aquí mismo veremos que R no es de factorización única.

5. Los enteros Gaussianos $\mathbb{Z}[i] \subseteq \mathbb{C}$, donde $i^2 = -1$, forman un anillo euclideano con $\delta(a + bi) = a^2 + b^2$ para $a + bi \in \mathbb{Z}[i]$ con $a, b \in \mathbb{Z}$, lo que veremos en esta misma sección.

6. El anillo $R = \{a + b(1 + \sqrt{-19})/2 \mid a, b \in \mathbb{Z}\}$ es principal sin ser euclideano, como veremos aquí mismo.

Proposición 2.30 *El anillo R del Ejemplo 4 no es de factorización única.*

Demostración: Para $x = a + b\sqrt{-5} \in R$, definimos $\overline{x} = a - b\sqrt{-5}$ y la **norma** de x como $N(x) = x\overline{x} = a^2 + 5b^2 \in \mathbb{Z}$.

Es fácil ver que $N(xy) = N(x)N(y)$ para todos $x, y \in R$.

Afirmamos que $R^\star = \{x \in R \mid N(x) = 1\} = \{\pm 1\}$, las unidades de R.

Si $N(x) = x\overline{x} = 1$, es claro que $x^{-1} = \overline{x} \in R$; y entonces $x \in R^\star$.

Recíprocamente, si $xz = 1$, entonces $N(x)N(z) = N(xz) = N(1) = 1$. Esto implica que $N(x) = \pm 1$; pero -1 no se puede escribir como $a^2 + 5b^2$ con $a, b \in \mathbb{Z}$. Así, $N(x) = 1$.

Ahora bien, $a^2 + 5b^2 = 1 \Rightarrow a = \pm 1, b = 0$.

En R, el número 3 es irreducible, pues $3 = ab$ con $a, b \in R$ implica que $9 = N(3) = N(a)N(b)$; y entonces $N(a) = 1, 3$ ó 9; pero también tenemos que $N(a) = 1 \Rightarrow a \in R^\star$, mientras que $N(a) = 9 \Rightarrow N(b) = 1 \Rightarrow b \in R^\star$. Por otra parte, 3 no se puede escribir como $c^2 + 5d^2$ con $c, d \in \mathbb{Z}$.

De manera similar puede verse que $2 + \sqrt{-5}$ es irreducible; y claramente no es asociado de 3.

La igualdad $9 = 3^2 = (2 + \sqrt{-5})(2 - \sqrt{-5})$ pone de manifiesto la falta de unicidad de factorizaciones en R. □

Observación. En todo anillo conmutativo es válida la identidad

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2;$$

que puede verificarse directamente. En el caso de los enteros, es natural considerar a los enteros Gaussianos $\mathbb{Z}[i]$:

Escribiendo $\alpha = a + bi$ con $a, b \in \mathbb{Z}$; tenemos la conjugación compleja $\overline{\alpha} = a - bi$, de manera que $\delta(\alpha) = \alpha\overline{\alpha} = a^2 + b^2$; y si $\beta = c + di$ con $c, d \in \mathbb{Z}$, entonces $\delta(\beta) = c^2 + d^2$. Resulta que $\alpha\beta = (ac - bd) + (ad + bc)i$; mientras que nuestra identidad dice que $\delta(\alpha\beta) = \delta(\alpha)\delta(\beta)$.

Lema 2.31 $\mathbb{Z}[i]^\star = \{\pm 1, \pm i\}$.

Demostración: Como $(a + bi)^{-1} = (a^2 + b^2)^{-1}(a - bi) \in \mathbb{C}$, el elemento $a + bi \in \mathbb{Z}[i]$ es unidad si y sólo si $\delta(a + bi) = a^2 + b^2 = 1$; y esto ocurre si y sólo si $a + bi \in \{\pm 1, \pm i\}$. □

Un anillo relacionado con los enteros Gaussianos y que usaremos, es $\mathbb{Q}[i] = \{a + bi \in \mathbb{C} \mid a, b \in \mathbb{Q}\}$, el cual es un campo, pues si $x = a + bi \neq 0$, entonces $x^{-1} = (a^2 + b^2)^{-1}(a - bi) \in \mathbb{Q}[i]$.

Teorema 2.32 *$\mathbb{Z}[i]$ es un anillo euclideano con $\delta(a + bi) = a^2 + b^2$.*

Demostración: Dados $\alpha, \beta \in \mathbb{Z}[i]$ con $\beta \neq 0$, se tiene que $\alpha\beta^{-1} = m + ni$ con $m, n \in \mathbb{Q}$.

Tenemos que existen $u, v \in \mathbb{Z}$ tales que

$$|m - u| \leq \frac{1}{2}, \ |n - v| \leq \frac{1}{2}.$$

Sean $c = m - u$ y $d = n - v$. Si escribimos $q = u + vi$ y $r = \beta(c + di)$, tendremos que $\alpha = \beta q + r$, de manera que $r = \alpha - \beta q \in \mathbb{Z}[i]$. Por otro lado,

$$\delta(r) = \delta(\beta)\delta(c + di) \leq \delta(\beta)(\frac{1}{4} + \frac{1}{4}) = \frac{1}{2}\delta(\beta).$$

Así, $\delta(r) < \delta(\beta)$. □

Lema 2.33 *Sea p un número primo tal que existan enteros x, y, z satisfaciendo $x^2 + y^2 = zp$ y también $(z, p) = 1$. Entonces existen enteros a, b tales que $a^2 + b^2 = p$.*

Demostración: Primero afirmamos que p no es irreducible en $\mathbb{Z}[i]$: Si suponemos p primo en $\mathbb{Z}[i]$, se tiene que $p \mid (x^2 + y^2) = (x + yi)(x - yi)$, lo cual implica que $p \mid (x + yi)$ ó bien que $p \mid (x - yi)$; por lo que en todo caso p divide a ambos. Entonces $p^2 \mid (x^2 + y^2) = zp$, que es una contradicción.

Así, $p = (a + bi)(c + di)$ con $a, b, c, d \in \mathbb{Z}$, $a^2 + b^2 \neq 1$ y $c^2 + d^2 \neq 1$. Entonces $p^2 = \delta(p) = (a^2 + b^2)(c^2 + d^2)$, por lo que $p = a^2 + b^2$. □

Lema 2.34 *Sea p un número primo tal que $p \equiv 1 \pmod 4$. Entonces existe un entero a tal que $a^2 \equiv -1 \pmod p$.*

Demostración: Sea

$$a = 1 \cdot 2 \cdot 3 \cdots \frac{p-1}{2} = (-1) \cdot (-2) \cdot (-3) \cdots (-\frac{p-1}{2}), \text{ entonces}$$

$$a^2 = 1 \cdot 2 \cdot 3 \cdots \frac{p-1}{2} \cdot (-\frac{p-1}{2}) \cdots (-3) \cdot (-2) \cdot (-1) \equiv (p-1)! \equiv -1 \pmod p.$$

Esta última congruencia es el Teorema de Wilson. □

Teorema 2.35 *Sea p un número primo.*

a) Si $p \equiv 1 \pmod 4$, entonces existen enteros a, b tales que $p = a^2 + b^2$.

b) Si $p \equiv 3 \pmod 4$, no existen enteros a, b con $p = a^2 + b^2$.

Demostración: *a)* Por el Lema 2.34, existen a, z tales que $a^2 + 1 = zp$; pero si suponemos que $|a| < (p/2)$, tendremos $(z, p) = 1$, pues $(p^2/4) + 1 < p^2$. La conclusión se obtiene aplicando el Lema 2.33.

b) Para todo entero n, se tiene que $(2n)^2 \equiv 0 \pmod 4$; mientras que $(2n + 1)^2 \equiv 1 \pmod 4$. De manera que si $a, b \in \mathbb{Z}$, es claro que se tiene $a^2 + b^2 \equiv 0, 1, 2 \pmod 4$. Así, $p \neq a^2 + b^2$. □

Si a es un entero y p es un número primo, decimos que a es un **residuo cuadrático** $(\text{mod } \mathbf{p})$ cuando la congruencia $x^2 \equiv a (\text{mod } p)$ tiene solución. En caso contrario, a es un **residuo no cuadrático**.

Definimos el **símbolo de Legendre** como sigue:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & \text{cuando } a \text{ es un residuo cuadrático,} \\ -1, & \text{cuando } a \text{ es un residuo no cuadrático.} \end{cases}$$

Lo anterior para $p \nmid a$; mientras que $\left(\frac{a}{p}\right) = 0$, si $p \mid a$.

Teorema 2.36 (Criterio de Euler) *Sean $a, b \in \mathbb{Z}$ y $p \neq 2$ un número primo. Entonces*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \quad y \quad \left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

Demostración: $(\mathbb{Z}/p\mathbb{Z})^\star$ es cíclico por el Teorema 2.15. Supongamos que el entero r representa a un generador de este grupo. Entonces existe un entero t tal que $a \equiv r^t \pmod{p}$; y a es un residuo cuadrático si y sólo si t es par. De aquí se obtiene la igualdad

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Por otro lado, $(\mathbb{Z}/p\mathbb{Z})^\star$ tiene un único subgrupo de orden 2 y por tanto un único elemento de orden 2, que es -1; de manera que tenemos la igualdad $r^{(p-1)/2} \equiv -1 \pmod{p}$.

Como r es un residuo no cuadrático, se tiene $\left(\frac{r}{p}\right) \equiv -1 \pmod{p}$.

La otra identidad deseada se obtiene de los siguientes cálculos:

$$\left(\frac{a}{p}\right) \equiv \left(\frac{r^t}{p}\right) \equiv (-1)^t \equiv (r^{(p-1)/2})^t \equiv (r^t)^{(p-1)/2} \equiv a^{(p-1)/2} \pmod{p}. \quad \square$$

Teorema 2.37 (Fermat) *Un entero positivo n es la suma de dos cuadrados enteros si y sólo si para todo primo p tal que $p \mid n$ y que $p \equiv 3 \pmod{4}$, el exponente de la máxima potencia de p que divida a n sea par.*

Demostración: La implicación $\Leftarrow$ es debida a $2 = 1^2 + 1^2$, a la identidad $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$; y al Teorema 2.35 *a*).

Recíprocamente, suponiendo que $n = a^2 + b^2$ con $a, b \in \mathbb{Z}$; y que p es un primo tal que $p \equiv 3 \pmod{4}$, $n = p^{2k+1}m$ y $p \nmid m$, exhibiremos una contradicción.

Sean $(a, b) = d$, $a = da_1$ y $b = db_1$, de manera que $(a_1, b_1) = 1$. Escribimos $n = d^2 n_1$, para tener $n_1 = a_1^2 + b_1^2$ con $p \mid n_1$ sin que p divida a los dos números a_1, b_1. Si es el caso de que $p \nmid a_1$, entonces existe c tal que $a_1 c \equiv b_1 \pmod{p}$ y se tiene $n_1 = a_1^2(1 + c^2) \equiv 0 \pmod{p}$. Esto implica la contradicción $c^2 \equiv -1 \pmod{p}$. El caso $p \nmid b_1$ es similar. $\square$

Proposición 2.38 *a) Todo factor primo de un entero de la forma $4m^2 + 1$ es de la forma $4n + 1$.*

b) El número de primos p con $p \equiv 1 \pmod{4}$ es infinito.

Demostración: *a*) $2 \nmid (4m^2+1)$. Si $p \equiv 3 \pmod 4$ y $p \mid (4m^2+1)$, entonces $\left(\frac{-1}{p}\right) = 1$, que es falso.

b) Supongamos que $\{p_1, ..., p_n\}$ es la lista completa de los primos p_i tales que $p_i \equiv 1 \pmod 4$, entonces $p_i \nmid (4p_1^2 \cdots p_n^2 + 1)$ para $1 \leq i \leq n$, por lo que el inciso anterior garantiza la existencia de otro primo $q \equiv 1 \pmod 4$ tal que $q \mid (4p_1^2 \cdots p_n^2 + 1)$. Esta contradicción concluye la demostración. □

Definiciones. (Motzkin) En un dominio D, definimos:

a) Un subconjunto P es un **ideal producto** cuando $P(D \setminus \{0\}) \subseteq P$.

b) Dado un subconjunto S de D, definimos su **conjunto derivado total** como $B = \{b \in D \mid \text{existe } a \in D \text{ con } a + bD \subseteq S\}$.

c) Dado un subconjunto S de D, su **conjunto derivado** es $S' = B \cap S$.

Observaciones. Las siguientes afirmaciones son claras:

a) Si S es un ideal producto, entonces S' también lo es.

b) $(S_1 \subseteq S) \Rightarrow (S_1' \subseteq S')$.

Teorema 2.39 (Motzkin) *a) Si (D, δ) es un anillo euclideano y $P_i = \{a \in D \mid \delta(a) \geq i\}$ para $i \in \mathbb{N}$, entonces cada P_i es un ideal producto, $\cap P_i = \varnothing$; y se satisfacen las relaciones $P_i' \subseteq P_{i+1}$ para todo $i \in \mathbb{N}$.*

b) Recíprocamente, dados un dominio D y una sucesión

$$D \setminus \{0\} = P_0 \supseteq P_1 \supseteq \cdots \tag{2.3}$$

de ideales producto con intersección vacía tales que $P_i' \subseteq P_{i+1}$ para todo $i \in \mathbb{N}$, entonces la función δ dada por $\delta(b) = i$ si $b \in (P_i \setminus P_{i+1})$ transforma a D en un anillo euclideano.

Demostración: *a*) Si D es euclideano, claramente todo P_i es un ideal producto y $\cap P_i = \varnothing$.

Sea $b \in P_i'$. Esto es debido a que existe $a \in D$ tal que para todo $q \in D$ se tiene que $(a - bq) \in P_i$, es decir, $\delta(a - bq) \geq i$. Como D es euclideano, esto implica que $\delta(b) \geq i + 1$. Así, $P_i' \subseteq P_{i+1}$.

b) Esto es claro. □

Corolario 2.40 *Dado un dominio D, hay una equivalencia entre algoritmos euclideanos en D y sucesiones* (2.3) *de ideales producto P_i con intersección vacía tales que $P_i' \subseteq P_{i+1}$ para todo $i \in \mathbb{N}$.*

Si para un dominio D, tenemos dos algoritmos euclideanos dados por sucesiones P_i y $\overline{P}_i$ tales que $P_i \subseteq \overline{P}_i$ para todo i, decimos que el algoritmo correspondiente a P_i es **más rápido**.

Corolario 2.41 *Dado un anillo euclideano D, siempre existe un algoritmo más rápido, que corresponde a la sucesión $D \setminus \{0\} = P_0 \supseteq P_0' \supseteq P_0'' \supseteq \cdots$.*

Corolario 2.42 *Un dominio D admite un algoritmo euclideano si y sólo si la sucesión $D \setminus \{0\} = P_0 \supseteq P_0' \supseteq P_0'' \supseteq \cdots$ tiene intersección vacía.*

Por el resto de la sección adoptamos la notación $P_0 = D \setminus \{0\}$.

Proposición 2.43 *Sea $D = \{a + b(1 + \sqrt{-19})/2 \mid a, b \in \mathbb{Z}\}$.*
a) $D^\star = \{\pm 1\}$.
b) $P_0' = P_0 \setminus D^\star$.
c) $P_0'' = P_0'$.
d) D no es un anillo euclideano.

Demostración: *a*) Usando la norma N de $\mathbb{Q}[\sqrt{-19}]$ se ve que

$$N(a + \frac{b}{2}(1 + \sqrt{-19})) = a^2 + ab + 5b^2 = 1 \Rightarrow a = \pm 1,\ b = 0.$$

b) Si $u \in D^\star$ con $uv = 1$ y $b \in D$, entonces $b + u(-vb) = 0$; que no está en P_0, por lo que $u \notin P_0'$.

Si $b \notin D^\star$, escribimos $a = 1$ para tener $a + bx \neq 0$ para todo $x \in D$; lo cual demuestra que si $b \notin D^\star$, entonces $b \in P_0'$.

c) A partir de la definición de "conjunto derivado total", vemos que al pasar de P_0' a P_0'', solamente se excluyen aquellos elementos $b \in P_0'$ tales que para todo $a \in D$ se tenga que $b \mid a$ ó bien $b \mid (a + u)$ con $u \in D^\star$; pero no hay tales elementos:

Tomamos $a = 2, (1+\sqrt{-19})/2$ y buscamos b que divida simultáneamente a uno de $\{1, 2, 3\}$ y a uno de $\{(\pm 1+\sqrt{-19})/2, (3+\sqrt{-19})/2\}$; pero entonces $N(b)$ tiene que dividir a uno de $\{1, 4, 9\}$ y también a uno de $\{5, 7\}$, que son los conjuntos de las normas permitidas. Esto implica que $N(b) = 1$, es decir, que b es una unidad, fuera de P_0'.

d) Acabamos de ver que la cadena $D \setminus \{0\} = P_0 \supseteq P_0' \supseteq P_0'' \supseteq \cdots$, se detiene en $P_0'' \neq \varnothing$. La conclusión se obtiene del Corolario 2.42. □

Un **anillo euclideano generalizado** es un dominio R provisto de una función $\delta : R \setminus \{0\} \to \mathbb{N}$ que satisface la siguiente condición:

Dados $a, b \in R \setminus \{0\}$, tales que $b \nmid a$, existen $c, d \in R$ con $\delta(ca - db) < \delta(b)$ y $ca - db \neq 0$.

Proposición 2.44 *Sea R un anillo euclideano generalizado, entonces R es principal.*

Demostración: Sean $I \neq (0)$ un ideal de R y $a \in I$ un elemento con $\delta(a)$ mínimo. Veamos que $I = (a)$:

Si $b \in I$, entonces $a \mid b$ ó bien existen $c, d \in R$ tales que $\delta(ca - db) < \delta(a)$; esto último es imposible porque $ca - db \in I$. Así, $a \mid b$. □

Si en un anillo euclideano generalizado R se cumple $\delta(ab) = \delta(a)\delta(b)$ para todos $a, b \in R \setminus \{0\}$, entonces podemos extender δ al campo de fracciones Q de R; y la condición que define a los anillos euclideanos generalizados puede ser reemplazada por la siguiente:

Para todo $e \in Q \setminus R$, existen $c, d \in R$ tales que $ce - d \neq 0,\ \delta(ce - d) < 1$.

Observación. Si $D = \{a + b(1 + \sqrt{-19})/2 \mid a, b \in \mathbb{Z}\}$, entonces

$$Q = \mathbb{Q}[\sqrt{-19}] \ (= \{a + b\sqrt{-19} \mid a, b \in \mathbb{Q}\} \ ).$$

Esto es porque claramente se tiene la inclusión $\supseteq$; y porque $\mathbb{Q}[\sqrt{-19}]$ es un campo:

$$a^2 + b^2 \neq 0 \Rightarrow (a + b\sqrt{-19})^{-1} = \frac{1}{a^2 + 19b^2}(a - b\sqrt{-19}).$$

Proposición 2.45 *$D = \{a + b(1 + \sqrt{-19})/2 \mid a, b \in \mathbb{Z}\}$ es un anillo euclideano generalizado.*

Demostración: Dado $e \in Q \setminus R$, escribimos $e = (a + b\sqrt{-19})/f$ con $a, b, f \in \mathbb{Z}$ y $(a, b, f) = 1$. Supongamos que $f \geq 5$.

Existen enteros x, y, z, q, r tales que $xa + yb + zf = 1$, $ay - 19bx = fq + r$, $|r| \leq f/2$.

Escribimos $c = y + x\sqrt{-19}$, $d = q - z\sqrt{-19}$, para tener

$$ce - d = \frac{(y + x\sqrt{-19})(a + b\sqrt{-19})}{f} - (q - z\sqrt{-19}) = \frac{r + \sqrt{-19}}{f}.$$

Este número tiene norma $0 \neq (r^2 + 19)/f^2 < 1$, porque $|r| \leq f/2$ y $f \geq 5$. Dejamos como ejercicio verificar los casos $f = 2, 3, 4$. □

Corolario 2.46 *$D = \{a + b(1 + \sqrt{-19})/2 \mid a, b \in \mathbb{Z}\}$ es un dominio principal que no admite estructura de anillo euclideano.*

Demostración: Esto es debido a las Proposiciones 2.43, 2.44 y 2.45. □

Ejercicios

1. Sea $m > 1$ un entero. Demuestre que todo entero positivo n se puede expresar de manera única como $n = c_s m^s + \cdots + c_1 m + c_0$, donde $0 \leq c_i < m$ para toda $0 \leq i \leq s$ y $c_s > 0$.

2. Determine los elementos irreducibles de $\mathbb{Z}[i]$.

3. Demuestre que existe un número infinito de primos $p \in \mathbb{Z}$ tales que $p \equiv 3 \pmod 4$.

4. Sea R un dominio principal. Demuestre que todo ideal primo de R distinto de (0) es máximo.

5. Sean R un dominio principal y $\mathfrak{a} \neq (0)$ un ideal. Demuestre que $R/\mathfrak{a}$ tiene un número finito de ideales.

6. Sean a y b enteros positivos tales que $a^2 = b^4 + b^3 + b^2 + b + 1$. Demuestre que $b = 3$.

7. Verifique los casos $f = 2, 3, 4$ de la Proposición 2.45.

2.7 Polinomios

Sean A un anillo conmutativo y T un símbolo nuevo, el **anillo de polinomios** en una variable T, escrito $A[T]$, consiste de todas las expresiones de la forma $a_0 + a_1T + \cdots + a_nT^n$, donde $n \in \mathbb{N}$, $a_i \in A$ para todo i. Cuando $a_n \neq 0$, se dice que el polinomio es de grado n.

Las operaciones que le dan estructura de anillo conmutativo a $A[T]$ son:

$$(a_0 + \cdots + a_nT^n) + (b_0 + \cdots + b_nT^n) = (a_0 + b_0) + \cdots + (a_n + b_n)T^n.$$

$$(a_0+\cdots+a_nT^n)(b_0+\cdots+b_mT^m) = c_0+\cdots+c_{n+m}T^{n+m} \text{ con } c_r = \sum_{i+j=r} a_ib_j$$

Una consecuencia inmediata de la definición de multiplicación de polinomios es que si A es un dominio, entonces $A[T]$ también lo es:

$$a_n, b_m \neq 0 \Rightarrow c_{n+m} = a_nb_m \neq 0 \Rightarrow (a_0+\cdots+a_nT^n)(b_0+\cdots+b_mT^m) \neq 0.$$

Si $g(T) = b_nT^n + \cdots + b_0$ con $b_n \neq 0$, se dice que b_n es el **coeficiente líder** de $g(T)$. Si $b_n = 1$, se dice que $g(T)$ es **mónico**.

Dados $f(T) = a_nT^n + \cdots + a_0$ y $g(T) = b_mT^m + \cdots + b_0$ con $g(T)$ mónico; si $m \leq n$, entonces $f - a_nT^{n-m}g$ es un polinomio de grado menor al de f. Este proceso puede continuar hasta obtener polinomios q, r tales que $f - qg = r$, donde $r = 0$ ó bien (grado r) $<$ (grado g). Este es el **algoritmo euclideano**.

Cuando k es un campo, el anillo $k[T]$ es euclideano con función $\delta =$ grado, pues $b_m \neq 0 \Rightarrow b_m$ es unidad, por lo que el primer paso del algoritmo es $f - b_m^{-1}a_nT^{n-m}g$; y los pasos sucesivos también son posibles. Así, $k[T]$ es un dominio principal y de factorización única.

Definimos inductivamente $A[X_1, ..., X_n] = A[X_1, ..., X_{n-1}][X_n]$.

El grado del **monomio** $X_1^{m_1} \cdots X_n^{m_n}$ es $m_1 + \cdots + m_n$. El grado de un polinomio $f \in A[X_1, ..., X_n]$ es el máximo grado de sus monomios. Un polinomio es **homogéneo** cuando todos sus monomios son del mismo grado.

Si $A[X_1, ..., X_n]_d$ es el conjunto de los polinomios homogéneos de grado d junto con cero, entonces

$$A[X_1, ..., X_n] = \bigoplus_{d \geq 0} A[X_1, ..., X_n]_d$$

es una suma directa de grupos abelianos aditivos.

Inductivamente, es claro que si A es un dominio, entonces $A[X_1, ..., X_n]$ también lo es; y que $\text{gr}(fg) = \text{gr}(f) + \text{gr}(g)$ para $f, g \in A[X_1, ..., X_n]$. De ahí se desprende que $A[X_1, ..., X_n]^\star = A^\star$. De manera que f y g son asociados si y sólo si existe $u \in A^\star$ tal que $g = uf$.

Cada $f \in A[X_1, ..., X_n]$ da origen a una función $f : A^n \to A$ dada por $(a_1, ..., a_n) \mapsto f(a_1, ..., a_n)$, resultado de substituir a_i en lugar de X_i.

Lema 2.47 *Sean k un campo, $a \in k$ y $f(X) \in k[X]$. Entonces*

$$(X-a) \mid f(X) \Leftrightarrow f(a) = 0.$$

Demostración: Por el algoritmo euclideano, existen $q(X), r(X) \in k[X]$ tales que $f(X) - (X-a)q(X) = r(X)$, donde $r(X)$ es cero o de grado cero. Como $f(a) = r(a) = r(X)$, se tiene la conclusión. □

Decimos que $a \in k$ es **raíz** de $f(X)$ de **multiplicidad** m cuando

$$(X-a)^m \mid f(X),\ (X-a)^{m+1} \nmid f(X).$$

Observación. Como $k[X]$ es de factorización única, es claro que $f(X)$ de grado n tiene cuando más n raíces, aún contando multiplicidades.

Teorema 2.48 (Fórmula de Interpolación de Lagrange) *Si $a_1, ..., a_n$ son n elementos distintos de un campo k mientras que $b_1, ..., b_n \in k$ son arbitrarios, entonces existe exactamente un polinomio $f(X) \in k[X]$ de grado no mayor que $n-1$ tal que $f(a_i) = b_i$ para $1 \leq i \leq n$.*

Demostración: El polinomio

$$f(X) = \sum_{j=1}^{n} b_j \frac{\prod_{i \neq j}(X - a_i)}{\prod_{i \neq j}(a_j - a_i)}$$

satisface los requisitos pedidos. Si $g(X)$ es otro polinomio que también los satisface, entonces $f(X) - g(X)$ es un polinomio de grado no mayor que $n-1$ con n raíces. Por tanto, $f(X) = g(X)$. □

Sea $f(X) = a_n X^n + \cdots + a_1 X + a_0 \in A[X]$, donde A es un dominio de factorización única. Definimos el **contenido** de $f(X)$ como

$$\mathfrak{c}(f) = \text{m.c.d.}\{a_0, a_1, ..., a_n\}.$$

Claramente, $\mathfrak{c}(f) \in A/A^\star$; pero normalmente elegimos un representante en A. Decimos que $f(X)$ es **primitivo** cuando $\mathfrak{c}(f) = 1$.

Lema 2.49 (Gauss) *Sea A un dominio de factorización única. Dos polinomios $f(X)$ y $g(X)$ son primitivos $\Leftrightarrow$ $f(X)g(X)$ es primitivo.*

Demostración: Si $f(X) = a_n X^n + \cdots + a_0$ y $g(X) = b_m X^m + \cdots + b_0$, se tiene $f(X)g(X) = \sum_i c_i X^i$ con $c_r = \sum_{i+j=r} a_i b_j$, de donde se ve que tanto $\mathfrak{c}(f)$ como $\mathfrak{c}(g)$ dividen a $\mathfrak{c}(fg)$.

Recíprocamente, si $f(X)g(X)$ no es primitivo, entonces existe un elemento irreducible $p \in A$ que divide a todos los coeficientes de $f(X)g(X)$. En el dominio $(A/(p))[X]$ se tiene $\overline{f}\,\overline{g} = \overline{fg} = 0$. Esto implica que $\overline{f} = 0$ o que $\overline{g} = 0$. Así, p divide a todos los coeficientes de $f(X)$ o a todos los coeficientes de $g(X)$. □

Lema 2.50 *Sean A un dominio de factorización única, k su campo de fracciones y $0 \neq f(X) \in k[X]$. Entonces podemos escribir $f(X) = cg(X)$ con $c \in k$ y $g(X) \in A[X]$ primitivo. Además, si $f(X) = c'h(X)$ con $c' \in k$ y $h(X) \in A[X]$ primitivo, entonces existen unidades $u, v \in A^\star$ satisfaciendo $c = uc'$, $g(X) = vh(X)$ y $uv = 1$.*

Demostración: Supongamos que $f(X) = e_n X^n + \cdots + e_1 X + e_0$. Como todo $e_i \in k$, existen $a_i, b_i \in A$ para $0 \leq i \leq n$ tales que $b_i \neq 0$ y $e_i = a_i/b_i$.

Sea $b = b_0 \cdots b_n$, entonces $bf(X) \in A[X]$ y existe $a \in A$ tal que $bf(X) = ag(X)$ con $g(X) \in A[X]$ primitivo; de manera que $f(X) = cg(X)$ con $c = a/b \in k$.

Si además, $f(X) = c'h(X)$ con $c' = a'/b' \in k$ y $h(X) \in A[X]$ primitivo, entonces $(a/b)g(X) = (a'/b')h(X)$; y por tanto, $b'ag(X) = ba'h(X)$; pero entonces $\mathfrak{c}(b'ag(X)) = \mathfrak{c}(ba'h(X))$. Esto significa que existe $u \in A^\star$ tal que $b'a = uba'$, es decir, $c = uc'$.

La igualdad $(a/b)g(X) = (a'/b')h(X)$ implica que también existe $v \in A^\star$ tal que $g(X) = vh(X)$ y $uv = 1$. □

Observación. En las condiciones del lema, si $f(X), g(X) \in A[X]$ son primitivos y existe $0 \neq a \in k$ tal que $f(X) = ag(X)$, entonces la unicidad demostrada implica que $a \in A^\star$.

Proposición 2.51 *Sea A un dominio de factorización única con campo de fracciones k y sea $p(X) \in A[X]$ primitivo. Entonces $p(X)$ se factoriza de manera única como producto de elementos irreducibles de $A[X]$, cuyos grados son los mismos que los provenientes de una factorización en $k[X]$.*

Demostración: Sea k el campo de fracciones de A. Entonces podemos escribir $p(X) = f_1(X) \cdots f_r(X)$ con cada $f_i(X)$ irreducible en $k[X]$.

El lema afirma que para cada i existen $a_i, b_i \in A$; $p_i(X) \in A[X]$ tales que $a_i b_i \neq 0$, con $p_i(X)$ primitivo y con

$$f_i(X) = \frac{a_i}{b_i} p_i(X).$$

Claramente, cada $p_i(X)$ es irreducible en $A[X]$.

De la igualdad

$$p(X) = f_1(X) \cdots f_r(X) = \frac{a_1 \cdots a_r}{b_1 \cdots b_r} p_1(X) \cdots p_r(X),$$

obtenemos $b_1 \cdots b_r p(X) = a_1 \cdots a_r p_1(X) \cdots p_r(X)$. Por el Lema de Gauss se concluye que $a_1 \cdots a_r = \mathfrak{c}(a_1 \cdots a_r p_1 \cdots p_r)$ y $b_1 \cdots b_r = \mathfrak{c}(b_1 \cdots b_r p)$ son asociados; y de ahí que

$$p(X) = p_1(X) \cdots p_r(X),$$

tal vez modificando un factor $p_i(X)$ multiplicativamente con una unidad de A, lo que demuestra la existencia de la factorización enunciada.

Para ver la unicidad, supongamos que $p(X) = h_1(X) \cdots h_s(X)$ con cada $h_j(X)$ irreducible en $A[X]$, entonces el Lema de Gauss garantiza que todo $h_j(X)$ es primitivo; y después el razonamiento que acabamos de ver demuestra que todo $h_j(X)$ es irreducible en $k[X]$. La unicidad de la factorización en $A[X]$ es consecuencia de la unicidad en $k[X]$ y del Lema 2.50. □

Corolario 2.52 *Si A es un dominio de factorización única, k su campo de fracciones y $f(X) \in A[X]$ es primitivo e irreducible, entonces también es irreducible en $k[X]$.*

Corolario 2.53 *Si $f(X) \in \mathbb{Z}[X]$ es mónico y admite factores de grado positivo en $\mathbb{Q}[X]$, entonces también los admite en $\mathbb{Z}[X]$.*

Teorema 2.54 *Si A es un dominio de factorización única, entonces $A[X]$ también lo es.*

Demostración: Dado $f(X) \in A[X]$, escribimos $f(X) = cg(X)$ con $c = \mathfrak{c}(f) \in A$ y $g(X) \in A[X]$ primitivo. La existencia y unicidad de la factorización de $f(X)$ en $A[X]$ se obtienen de las de c en A y $g(X)$ en $A[X]$. □

Corolario 2.55 *Sea A es un dominio de factorización única; entonces $A[X_1, ..., X_n]$ también lo es. En particular, para todo campo k y todo $n \in \mathbb{N}$, se tiene que $k[X_1, ..., X_n]$ es un dominio de factorización única.*

Lema 2.56 *Si $f(X) = a_r X^r + \cdots + a_1 X + a_0 \in \mathbb{Z}[X]$ y $m, n \in \mathbb{Z}$ son tales que $(X - \left(\frac{m}{n}\right)) \mid f(X)$ en $\mathbb{Q}[X]$ con $(m, n) = 1$, entonces $m \mid a_0$ y $n \mid a_r$.*

Demostración: Multiplicamos la igualdad $f\left(\frac{m}{n}\right) = 0$ por n^r para tener

$$a_r m^r + a_{r-1} m^{r-1} n + \cdots + a_1 m n^{r-1} + a_0 n^r = 0,$$

de donde se obtiene la conclusión. □

Usando el resultado anterior, se ve que los polinomios $X^2 - 2$, $X^2 - 3$ y $X^3 - 2$ son irreducibles en $\mathbb{Z}[X]$ y en $\mathbb{Q}[X]$. Esto implica que $\sqrt{2}$, $\sqrt{3}$ y $\sqrt[3]{2}$ son irracionales.

Teorema 2.57 (Criterio de Irreducibilidad de Eisenstein) *Dado un polinomio $f(X) = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{Z}[X]$ tal que existe un número primo p que satisface*

- $p \mid a_0$, $p \mid a_1, ...,$ $p \mid a_{n-1}$,
- $p \nmid a_n$,
- $p^2 \nmid a_0$,

se tiene que $f(X)$ es irreducible en $\mathbb{Q}[X]$.

Demostración: Dividiendo entre $\mathfrak{c}(f)$, se ve que es suficiente considerar el caso en que $f(X)$ es primitivo, lo que desde ahora suponemos.

Supongamos que $f(X) = g(X)h(X)$ con factores $g(X), h(X) \in \mathbb{Z}[X]$ no constantes. Si obtenemos una contradicción, habremos terminado en vista del Corolario 2.53.

Sea $\varphi : \mathbb{Z}[X] \to (\mathbb{Z}/p\mathbb{Z})[X]$ el morfismo natural, resultado de considerar los coeficientes de los polinomios $(\mathrm{mod}\, p)$. Aplicándolo a $f(X)$ tenemos que $\varphi(f(X)) = \overline{f}(X) = \overline{a}_n X^n = \overline{g}(X)\overline{h}(X)$.

Como $(\mathbb{Z}/p\mathbb{Z})[X]$ es de factorización única, se ve que $\overline{g}(X) = \overline{b}_m X^m$ y que $\overline{h}(X) = \overline{c}_s X^s$, con $m + s = n$ y con $b_m, c_s \in \mathbb{Z}$.

Si alguno de los números m ó s es cero, ya terminamos; si no, p divide a los términos constantes de g y h; pero entonces $p^2 \mid a_0$, que es una contradicción. □

Corolario 2.58 *Si p es un número primo, entonces $X^{p-1} + \cdots + X + 1$ es irreducible en $\mathbb{Q}[X]$.*

Demostración: Con la substitución $X = Y + 1$ se tiene que

$$X^{p-1} + \cdots + X + 1 = \frac{X^p - 1}{X - 1} = \frac{(Y+1)^p - 1}{Y}$$

$$= Y^{p-1} + \binom{p}{1} Y^{p-2} + \cdots + \binom{p}{p-1}$$

es irreducible por el teorema, ya que $p \mid \binom{p}{i}$ para $1 \leq i \leq p-1$, mientras que $p^2 \nmid \binom{p}{p-1} = p$. □

Ejemplo. El polinomio $f(X) = X^4 + 8X^3 + X^2 + 2X + 5$ es irreducible en $\mathbb{Q}[X]$.

1. Como $f(X)$ es primitivo, es suficiente demostrar su irreducibilidad en $\mathbb{Z}[X]$.

2. $f \equiv X^4 + X^2 + 1 (\mathrm{mod}\, 2)$, que no tiene raíces en $\mathbb{Z}/2\mathbb{Z}$; y por tanto no admite factores lineales en $(\mathbb{Z}/2\mathbb{Z})[X]$.

3. $f \equiv X(X^3 + 3X^2 + X + 2) (\mathrm{mod}\, 5)$. El polinomio $X^3 + 3X^2 + X + 2$ es irreducible en $(\mathbb{Z}/5\mathbb{Z})[X]$ al no tener raíces en $\mathbb{Z}/5\mathbb{Z}$.

4. De la incompatibilidad de las factorizaciones anteriores se obtiene la irreducibilidad de $f(X)$ en $\mathbb{Z}[X]$.

Si A es un anillo conmutativo y $A[X]$ es un anillo de polinomios en una variable, entonces existe una única función A-lineal $D : A[X] \to A[X]$ tal que $D(X^n) = nX^{n-1}$ para todo $n \in \mathbb{N}$. Decimos que $D(f) = f'$ es la **derivada** de f.

Como se verifica inmediatamente que $(X^r X^s)' = (r+s)X^{r+s-1} = (X^r)'X^s + X^r(X^s)'$, tenemos que $(fg)' = f'g + fg'$ para todos $f, g \in A[X]$.

Proposición 2.59 *Sean k un campo, $f(X) \in k[X]$, $a \in k$. Entonces a es raíz múltiple de $f(X)$ si y sólo si $f(a) = f'(a) = 0$.*

Demostración: Si $f(X) = (X-a)^2 g(X)$, entonces

$$f'(X) = (X-a)^2 g'(X) + 2(X-a)g(X),$$

por lo que $f'(a) = 0$.

Recíprocamente, si $f(X) = (X-a)g(X)$ con $g(a) \neq 0$, entonces $f'(X) = (X-a)g'(X) + g(X)$, por lo que $f'(a) = g(a) \neq 0$. □

Ejercicios

1. Sea D un dominio tal que $D[X]$ es principal. Demuestre que D es un campo.

2. Encuentre el número de monomios en n variables de grado d.

3. Sean R un anillo conmutativo y $f(X) \in R[X]$ un divisor de cero. Demuestre que existe $0 \neq a \in R$ tal que $af(X) = 0$.

4. Demuestre que $X^4 + 1,\ X^5 - X^2 + 1$ son irreducibles en $\mathbb{Q}[X]$.

5. Sean R un anillo conmutativo y $f(X) = a_n X^n + \cdots + a_1 X + a_0 \in R[X]$. Demuestre que $f(X) \in R[X]^\star \Leftrightarrow (a_0 \in R^\star$ y a_i es nilpotente para $1 \leq i \leq n)$.

6. Sean k un campo y A el subanillo de $k[X]$ de los polinomios de la forma $a_0 + a_2 X^2 + \cdots + a_n X^n$, es decir, sin término lineal. Exhiba un ideal no principal de A y demuestre que A no es de factorización única.

7. Sea $f(X) = a_{2n+1} X^{2n+1} + \cdots + a_1 X + a_0 \in \mathbb{Z}[X]$ primitivo tal que existe un número primo p con

   - $p^3 \nmid a_0$,
   - $p^2 \mid a_0,\ p^2 \mid a_1, ...,\ p^2 \mid a_n$,
   - $p \mid a_{n+1}, ...,\ p \mid a_{2n}$,
   - $p \nmid a_{2n+1}$,

   Demuestre que $f(X)$ es irreducible.

8. Sean $\mathbb{F}_p$ el campo de los enteros módulo p y $f : \mathbb{F}_p \to \mathbb{F}_p$ una función. Demuestre que f es polinomial.

9. Demuestre que para $0 \neq p(X_1, ..., X_n) \in k[X_1, ..., X_n]$, con k un campo infinito, siempre existen $a_1, ..., a_n \in k$ con $p(a_1, ..., a_n) \neq 0$.

2.8 Polinomios Simétricos, Resultante y Discriminante

Sea R un anillo conmutativo. Escribimos $\mathrm{Aut}_R(R[X_1,...,X_n])$ para representar al grupo de automorfismos σ del anillo de polinomios $R[X_1,...,X_n]$ tales que $\sigma(a)=a$ para todo $a\in R$.

Sea S_n el grupo simétrico. Este grupo actúa de manera natural en el anillo $R[X_1,...,X_n]$, es decir, existe un morfismo de grupos

$$\psi : S_n \to \mathrm{Aut}_R(R[X_1,...,X_n])$$

tal que $\psi(\alpha)(X_i)=X_{\alpha(i)}$ para todos $1\leq i\leq n,\ \alpha\in S_n$. Por brevedad, escribiremos α en lugar de $\psi(\alpha)$.

Consistentemente con la notación usada en Teoría de Grupos, escribimos $R[X_1,...,X_n]^{S_n}=\{f\in R[X_1,...,X_n]\mid \alpha(f)=f,\ \forall\ \alpha\in S_n\}$. Este es un subanillo de $R[X_1,...,X_n]$, cuyos elementos se llaman **polinomios simétricos**.

Los **polinomios simétricos elementales** son los siguientes:

$$\sigma_1=\sum_i X_i,\ \sigma_2=\sum_{i<j} X_iX_j,...,\ \sigma_n=X_1\cdots X_n.$$

Observemos que cada σ_i es homogéneo de grado i.

Sean A un subanillo de B y $\{b_1,...,b_n\}\subseteq B$. Entonces:

- Al anillo generado por $A\cup\{b_1,...,b_n\}$ lo escribimos $A[b_1,...,b_n]$.
- El conjunto $\{b_1,...,b_n\}$ es **algebraicamente independiente** sobre A cuando no existe $0\neq f\in A[X_1,...,X_n]$ tal que $f(b_1,...,b_n)=0$.
- Existe un morfismo de anillos $\theta : A[X_1,...,X_n]\to A[b_1,...,b_n]$ tal que $\theta(a)=a$ para todo $a\in A$ y $\theta(X_i)=b_i$ para todo $1\leq i\leq n$. Decir que $\{b_1,...,b_n\}$ es algebraicamente independiente sobre A es equivalente a decir que $\ker\theta=(0)$.

Teorema 2.60 *Sean A un anillo conmutativo y $\sigma_1,...,\sigma_n\in A[X_1,...,X_n]$ los polinomios simétricos elementales. Entonces:*

a) $A[\sigma_1,...,\sigma_n]=A[X_1,...,X_n]^{S_n}$, es decir, todo polinomio simétrico se puede expresar como un polinomio en los polinomios simétricos elementales.

b) $\{\sigma_1,...,\sigma_n\}$ es algebraicamente independiente sobre A, es decir, toda expresión en a) es única.

Demostración: La inclusión $A[\sigma_1,...,\sigma_n]\subseteq A[X_1,...,X_n]^{S_n}$ es clara. Para ver la recíproca procedemos como sigue.

Ordenamos los monomios de $A[X_1,...,X_n]$ lexicográficamente, es decir, escribimos $X_1^{a_1}\cdots X_n^{a_n} > X_1^{b_1}\cdots X_n^{b_n}$ cuando exista $1\leq i\leq n$ tal que $a_1=b_1,...,a_{i-1}=b_{i-1},\ a_i>b_i$.

Si $f(X)$ es un polinomio simétrico, entonces cada una de sus componentes homogéneas también lo es; por lo que suponemos que $f(X)$ es homogéneo de grado d.

Junto con cada monomio $M = aX_1^{m_1} \cdots X_n^{m_n}$ que aparezca en $f(X)$, aparecen también todos los monomios en su órbita, es decir, los que se obtienen a partir de M al permutar los exponentes, por lo que podemos suponer que tenemos un monomio que satisface $m_1 \geq \cdots \geq m_n$. Esto significa que M es máximo entre los que pertenecen a su órbita.

Supongamos que $aX_1^{r_1} \cdots X_n^{r_n}$ es el monomio máximo de f. Acabamos de ver que $r_1 \geq \cdots \geq r_n$. Observemos que $a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3} \cdots \sigma_n^{r_n}$ es un monomio en las σ_i, cuyo grado en las X_i es $(r_1-r_2)+2(r_2-r_3)+\cdots+nr_n = r_1+r_2+\cdots+r_n$; y que además es homogéneo en las X_i. Más aún, el monomio máximo en las X_i de $a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3} \cdots \sigma_n^{r_n}$ es $aX_1^{r_1} \cdots X_n^{r_n}$.

La conclusión es que $f - a\sigma_1^{r_1-r_2}\sigma_2^{r_2-r_3} \cdots \sigma_n^{r_n}$ es simétrico; y si no es cero, es homogéneo de grado d con monomio máximo menor que el de f.

Obtenemos a) por inducción en el orden lexicográfico al observar que el número de monomios en n variables de grado d es finito.

b) Sea $0 \neq p(T_1, ..., T_n) \in A[T_1, ..., T_n]$ un polinomio, entonces dado un monomio M_1 de $p(T_1, ..., T_n)$, siempre es posible escribirlo como $M_1 = aT_1^{r_1-r_2}T_2^{r_2-r_3} \cdots T_n^{r_n}$ con $r_i \in \mathbb{N}$ y $r_1 \geq \cdots \geq r_n$.

Consideremos al conjunto de los vectores $(r_1, ..., r_n) \in \mathbb{N}^n$ así obtenidos; y tomemos al máximo de ellos en el orden lexicográfico: $(s_1, ..., s_n)$. Resulta que el monomio máximo de $p(\sigma_1, ..., \sigma_n)$ en las X_i es $aX_1^{s_1} \cdots X_n^{s_n}$, que no se cancela con ningún otro. Se tiene pues que $p(\sigma_1, ..., \sigma_n) \neq 0$. □

Consideremos ahora el siguiente problema natural: Dados dos polinomios f y g en $k[X]$ con k un campo, ¿existe algún criterio para determinar si estos polinomios tienen un factor común no constante? El siguiente resultado es una respuesta positiva.

Teorema 2.61 (Sylvester) *Sean k un campo, $f(X) = a_0X^m + \cdots + a_m$ y $g(X) = b_0X^n + \cdots + b_n$ polinomios en $k[X]$ con $a_0 \neq 0 \neq b_0$ y $R(f, g)$ el siguiente determinante $(m + n) \times (m + n)$:*

$$\begin{vmatrix}
a_0 & a_1 & \cdots & a_m & & & & \\
 & a_0 & a_1 & \cdots & a_m & & & \\
 & & \ddots & & & \ddots & & \\
 & & & a_0 & a_1 & \cdots & a_m & \\
b_0 & b_1 & \cdots & b_n & & & & \\
 & b_0 & b_1 & \cdots & b_n & & & \\
 & & \ddots & & & \ddots & & \\
 & & & b_0 & b_1 & \cdots & b_n &
\end{vmatrix}$$

Entonces f y g tienen un factor común no constante si y sólo si $R(f, g) = 0$.

Demostración: Primero afirmamos que f y g tienen un factor común no constante p si y sólo si existen polinomios h y q tales que $fh = gq$, con $h \neq 0 \neq q,\ \operatorname{gr} h < \operatorname{gr} g = n,\ \operatorname{gr} q < \operatorname{gr} f = m$.

Si $f = p\alpha, g = p\beta$ con $\operatorname{gr} p \geq 1$, entonces $f\beta = p\alpha\beta = g\alpha$, por lo que $h = \beta$ y $q = \alpha$ funcionan. El recíproco es claro porque $k[X]$ es un dominio de factorización única.

Escribimos los polinomios $h(X) = c_0X^{n-1} + c_1X^{n-2} + \cdots + c_{n-1}$ y $-q(X) = d_0X^{m-1} + d_1X^{m-2} + \cdots + d_{m-1}$; y tratamos de resolver $fh = gq$ para $c_i, d_j \in k$.

Comparando los coeficientes de las distintas X^i en fh y en gq, tenemos que nuestro problema se reduce a resolver el siguiente sistema de ecuaciones lineales en las incógnitas c_i y d_j:

$$\begin{array}{lllllll}
a_0c_0 & & & = & -\ b_0\ d_0 & & \\
a_1c_0 & +\ a_0c_1 & & = & -\ b_1\ d_0 & -\ b_0d_1 & \\
a_2c_0 & +\ a_1c_1 & +\ a_0\ c_2 & = & -\ b_2\ d_0 & -\ b_1d_1 & -\ b_0d_2 \\
 & & \cdots & = & \cdots & & \\
 & & a_mc_{n-1} & = & -b_nd_{m-1} & &
\end{array}$$

Este sistema puede escribirse como

$$M \begin{pmatrix} c_0 \\ \vdots \\ c_{n-1} \\ d_0 \\ \vdots \\ d_{m-1} \end{pmatrix} = 0,$$

donde M es una matriz $(m+n) \times (m+n)$ tal que al determinante de su transpuesta le llamamos $R(f,g)$. $\square$

En las condiciones del teorema, el determinante $R(f,g)$ es el **resultante** de f y g.

Sean k un campo y $A = k[X_1, ..., X_m, Y_1, ..., Y_n]$ el anillo de polinomios en las $m+n$ variables indicadas. Consideremos los polinomios en $A[T]$:

$$f(T) = a_0 \prod_{i=1}^{m} (T - X_i) = a_0T^m + a_1T^{m-1} + \cdots + a_m,$$

$$g(T) = b_0 \prod_{i=1}^{n} (T - Y_i) = b_0T^n + b_1T^{n-1} + \cdots + b_n;$$

donde $a_0, b_0 \in k^\star$. Escribimos

$$\sigma_i = \sum_{r_1 < \cdots < r_i} X_{r_1} \cdots X_{r_i},\ \tau_j = \sum_{s_1 < \cdots < s_j} Y_{s_1} \cdots Y_{s_j};$$

de manera que $a_i, b_j, \sigma_i, \tau_j \in A$ para $0 \leq i \leq m$ y $0 \leq j \leq n$ con

$$a_i = (-1)^i a_0 \sigma_i, \quad b_j = (-1)^j b_0 \tau_j, \text{ para } 1 \leq i \leq m,\ 1 \leq j \leq n. \tag{2.4}$$

El **resultante genérico** $R(f,g)$ de f y g es el siguiente determinante $(m+n) \times (m+n)$:

$$\begin{vmatrix}
a_0 & a_1 & \cdots & a_m & & & & \\
 & a_0 & a_1 & \cdots & a_m & & & \\
 & & \ddots & & & \ddots & & \\
 & & & a_0 & a_1 & \cdots & a_m & \\
b_0 & b_1 & \cdots & b_n & & & & \\
 & b_0 & b_1 & \cdots & b_n & & & \\
 & & \ddots & & & \ddots & & \\
 & & & b_0 & b_1 & \cdots & b_n &
\end{vmatrix}$$

mientras que el **discriminante genérico** de f es el polinomio

$$D = \prod_{i<j} (X_i - X_j)^2.$$

Aplicando dos veces el Teorema 2.60 *a*) para la acción natural del grupo $G = S_m \times S_n$; y considerando la ecuación (2.4), tenemos que

$$A^G = k[\sigma_1, ..., \sigma_m, \tau_1, ..., \tau_n] = k[a_1, ..., a_m, b_1, ..., b_n].$$

Observación. Notemos que $R(f,g) \in A^G$; y que A^G es un anillo de polinomios en las variables σ_i, τ_j.

Teorema 2.62 $R(f,g) = a_0^n b_0^m \prod_{i,j}(X_i - Y_j)$.

Demostración: Usando las expresiones 2.4, vemos que $R(f,g)$ es el producto de $a_0^n b_0^m$ por un polinomio en σ_i, τ_j.

Fijamos i, j y llamamos K al campo de fracciones de A. Aplicando el Teorema de Sylvester a $f(T), g(T) \in K[T]$, obtenemos lo siguiente: Al considerar a $R(f,g)$ como polinomio en X_i y evaluarlo en Y_j, el resultado es cero, ya que en esas condiciones f y g tienen un factor común no trivial.

Como $X_i - Y_j$ es mónico, existen $q, r \in A$ tales que X_i no aparece en r y $R(f,g) = (X_i - Y_j)q + r$. El párrafo anterior garantiza que $r = 0$, por lo que obtenemos $(X_i - Y_j) \mid R(f,g)$ en A.

Escribiendo $S = a_0^n b_0^m \prod_{i,j}(X_i - Y_j)$, tenemos que $S \mid R(f,g)$ en A, porque todo $X_i - Y_j$ es irreducible; pero $S, R(f,g) \in A^G$. Esto implica que $S \mid R(f,g)$ en A^G.

Por un lado, el término de grado máximo en τ_n que aparece en $R(f,g)$ es $(-1)^{mn} a_0^n b_0^m \tau_n^m$. Por el otro lado, observando que

$$\prod_{i=1}^{m} g(X_i) = b_0^m \prod_{i,j}(X_i - Y_j),$$

se tiene que

$$S = a_0^n \prod_{i=1}^{m} g(X_i) = a_0^n \prod_{i=1}^{m} (b_0 X_i^n + b_1 X_i^{n-1} + \cdots + b_n); \qquad (2.5)$$

donde se ve que el término de grado máximo en τ_n que aparece en S es $(-1)^{mn} a_0^n b_0^m \tau_n^m$. Concluimos que $R(f,g) = S$. $\square$

El caso $g(T) = f'(T)$ con $f(T) = a_0 \prod_{i=1}^{n}(T - X_i)$ da lugar al siguiente resultado:

Teorema 2.63 $R(f, f') = (-1)^{n(n-1)/2} a_0^{2n-1} D$.

Demostración: A partir de $f(T) = a_0 \prod_{i=1}^{n}(T - X_i)$, tenemos que

$$f'(T) = a_0 \sum_{i=1}^{n} (T - X_1) \cdots \widehat{(T - X_i)} \cdots (T - X_n); \text{ y también}$$

$$f'(X_i) = a_0 (X_i - X_1) \cdots (X_i - X_{i-1})(X_i - X_{i+1}) \cdots (X_i - X_n).$$

De la proposición anterior y de (2.5) se obtiene

$$R(f, f') = a_0^{n-1} \prod_{i=1}^{n} f'(X_i) =$$

$$(-1)^{n(n-1)/2} a_0^{2n-1} \prod_{i<j} (X_i - X_j)^2 = (-1)^{n(n-1)/2} a_0^{2n-1} D. \ \square$$

Si k es un campo y $f(T) = \prod_{i=1}^{n}(T - r_i) \in k[X]$ con todo $r_i \in k$, entonces el **discriminante** de f es

$$\prod_{i<j} (r_i - r_j)^2.$$

Observación. Un polinomio $f(T) \in k[X]$ tiene raíces múltiples si y sólo si su discriminante se anula; y esto ocurre si y sólo si $f(T)$ y $f'(T)$ tienen un factor común de grado positivo.

Ejercicios

1. Calcule el discriminante de $X^2 + bX + c$ y el de $X^3 + pX + q$.
2. Exprese $X_1^2 + X_2^2 + X_3^2$ y $X_1^3 + X_2^3 + X_3^3$ como polinomios en $\sigma_1, \sigma_2, \sigma_3$.
3. Demuestre que el resultante de $a_0 X^2 + a_1 X + a_2$ y $b_0 X^2 + b_1 X + b_2$ es $a_2^2 b_0^2 - a_1 a_2 b_0 b_1 + a_0 a_2 b_1^2 + a_1^2 b_0 b_2 - 2a_0 a_2 b_0 b_2 - a_0 a_1 b_1 b_2 + a_0^2 b_2^2$.
4. Demuestre que el discriminante de $X^5 + pX + q$ es $2^8 p^5 + 5^5 q^4$; y que el de $X^7 + pX + q$ es $-2^6 3^6 p^7 - 7^7 q^6$.
5. Demuestre que el discriminante del polinomio $X^3 - a_1 X^2 + a_2 X - a_3$ es $-4a_1^3 a_3 + a_1^2 a_2^2 + 18 a_1 a_2 a_3 - 4a_2^3 - 27 a_3^2$.

2.9 Módulos y Anillos Noetherianos

Dado un anillo R, se dice que M es un R-**módulo (izquierdo)** cuando M es un grupo abeliano ante una operación $+$ que posee además una multiplicación $R \times M \to M$ dada por $(r, m) \mapsto rm$ tal que

- $1m = m$ para todo $m \in M$.
- $(a + b)m = am + bm$ para todos $a, b \in R,\ m \in M$.
- $(ab)m = a(bm)$ para todos $a, b \in R,\ m \in M$.
- $a(m + n) = am + an$ para todos $a \in R,\ m, n \in M$.

Proponemos como ejercicio informal, definir los conceptos de submódulo y de morfismo de R-módulos; así como el de R-módulo derecho.

Ejemplos. Algunos R-módulos importantes son:

1. Cuando R es un campo, un R-módulo es lo mismo que un espacio vectorial.
2. Todo grupo abeliano A es un $\mathbb{Z}$-módulo de manera natural: na significa $a + \cdots + a$, con n sumandos para $n \in \mathbb{N}$ y $a \in A$. También $(-1)a = -a$.
3. Todo ideal izquierdo de R es un R-módulo izquierdo.
4. Si $\mathfrak{a}$ es un ideal izquierdo de R, entonces $R/\mathfrak{a}$ también es un R-módulo izquierdo ante $r(x + \mathfrak{a}) = rx + \mathfrak{a}$, para $r, x \in R$.
5. Si M es un R-módulo izquierdo y N es un submódulo, entonces M/N también es un R-módulo izquierdo ante $r(x + N) = rx + N$, para $r \in R,\ x \in M$.

Por brevedad, diremos R-módulo en lugar de R-módulo izquierdo.

Un R-módulo M es **irreducible** cuando sus únicos submódulos son (0) y M. Para todo R-módulo M, el conjunto $\operatorname{End}_R M$ de endomorfismos de M admite una estructura de anillo ante la suma de funciones y la multiplicación dada por composición de funciones.

Teorema 2.64 (Lema de Schur) *Sea M un R-módulo irreducible. Entonces $\operatorname{End}_R M$ es un anillo de división.*

Demostración: Sea $\psi : M \to M$ un morfismo no trivial. Entonces su imagen es M por tener que ser un submódulo de M. Por otro lado, su núcleo es (0), por la misma razón. Sabemos que la función inversa ψ^{-1} es un morfismo de grupos abelianos; y es fácil ver que también es un morfismo de R-módulos. □

Proposición 2.65 *Sea M un R-módulo. Las siguientes condiciones son equivalentes:*

1. *Todo submódulo es finitamente generado.*
2. *Toda cadena de submódulos estrictamente ascendente es finita.*
3. *Toda colección no vacía de submódulos tiene un máximo*

Demostración: 1) $\Rightarrow$ 2): Si $M_1 \subset M_2 \subset \cdots$; escribimos $N = \cup_i M_i$, que está finitamente generado, por lo que existe n tal que M_n contiene a esos generadores; y entonces la cadena termina en n.

2) $\Rightarrow$ 3): Si $\Sigma \neq \varnothing$ es una colección no vacía de submódulos en la que no existen máximos, entonces cualquier $N_1 \in \Sigma$ no es máximo y existe $N_2 \in \Sigma$ con $N_1 \subset N_2$. Procedemos inductivamente a partir de $N_1 \subset N_2 \subset \cdots \subset N_i$, ya que N_i no es máximo, existe $N_{i+1} \in \Sigma$ con $N_i \subset N_{i+1}$. Así, es posible construir una cadena ascendente infinita.

3) $\Rightarrow$ 1): Dado N submódulo de M, sea Σ la colección de los submódulos de N que son finitamente generados. Si $(m_1, ..., m_r)$ es un elemento máximo de Σ, entonces claramente $N = (m_1, ..., m_r)$. □

Un módulo es **Noetheriano** cuando satisface las condiciones de la proposición anterior. Un anillo R es **Noetheriano** cuando lo es como R-módulo.

Proposición 2.66 *Sea M un R-módulo Noetheriano. Entonces toda imagen homomorfa y todo submódulo de M son Noetherianos.*

Demostración: La afirmación sobre submódulos es clara. Para ver la otra, supongamos que $\overline{M}_1 \subset \overline{M}_2 \subset \cdots$ es una cadena ascendente de submódulos de M/N; y que $M_i = f^{-1}(\overline{M}_i)$ para cada i, donde $f : M \to M/N$ es el morfismo natural. Entonces la cadena ascendente $M_1 \subset M_2 \subset \cdots$ es finita, por lo que $\overline{M}_1 \subset \overline{M}_2 \subset \cdots$ también lo es. □

Proposición 2.67 *Sean M un R-módulo y N un submódulo tales que N y M/N son Noetherianos. Entonces M es Noetheriano.*

Demostración: Observemos que si $L_1 \subseteq L_2$ son submódulos de M tales que $(L_1 \cap N) = (L_2 \cap N)$ y $(L_1 + N)/N = (L_2 + N)/N$, entonces $L_1 = L_2$: Pues $m \in L_2 \Rightarrow$ existen $u, v \in N$, $m' \in L_1$ tales que $m' + u = m + v$, de manera que $m - m' = u - v \in (L_2 \cap N) = (L_1 \cap N)$; y por tanto $m = m' + (u - v) \in L_1$.

Si ahora $M_1 \subseteq M_2 \subseteq \cdots$ es una cadena ascendente de submódulos de M, entonces las cadenas ascendentes $(M_1 \cap N) \subseteq (M_2 \cap N) \subseteq \cdots$ y $(M_1 + N)/N \subseteq (M_2 + N)/N \subseteq \cdots$ se estabilizan en algún punto. La observación previa implica que la cadena $M_1 \subseteq M_2 \subseteq \cdots$ también se estabiliza en ese punto. □

Si $M_1, ..., M_n$ son R-módulos, definimos la **suma directa de módulos** $M = M_1 \oplus \cdots \oplus M_n$ como la suma directa de grupos abelianos con multiplicación $r(m_1, ..., m_n) = (rm_1, ..., rm_n)$ para todos $r \in R,\ m_i \in M_i$.

Observaciones. Las siguientes afirmaciones son inmediatas:

1. Si $M_1, ..., M_n$ son submódulos de M, entonces $M = M_1 \oplus \cdots \oplus M_n$ si y sólo si $M = M_1 + \cdots + M_n$ y para todo $1 \leq i \leq n$ se tiene que $(M_1 + \cdots + M_{i-1}) \cap M_i = (0)$.

2. Un R-módulo M es finitamente generado si y sólo si existe un morfismo suprayectivo de R-módulos $\psi : R \oplus \cdots \oplus R \to M$.

Proposición 2.68 *a) Toda suma directa finita de módulos Noetherianos es Noetheriana.*
b) Toda suma finita de módulos Noetherianos es Noetheriana.
c) Sean R un anillo Noetheriano y M un R-módulo finitamente generado, entonces M es Noetheriano.

Demostración: *a)* Sea $M = M_1 \oplus \cdots \oplus M_n$ con todo M_i Noetheriano. Procedemos por inducción en n: Tenemos que $M/M_n \cong M_1 \oplus \cdots \oplus M_{n-1}$ con M_n y $M_1 \oplus \cdots \oplus M_{n-1}$ Noetherianos. La conclusión se obtiene de la proposición anterior.
b) Si ahora $M = M_1 + \cdots + M_n$ con todo M_i Noetheriano, entonces M es imagen homomorfa de $M_1 \oplus \cdots \oplus M_n$.
c) Aquí, M es imagen homomorfa de $R \oplus \cdots \oplus R$. □

Proposición 2.69 *Sea $f : A \to B$ un morfismo suprayectivo de anillos con A Noetheriano. Entonces B es Noetheriano.*

Demostración: Si $\mathfrak{b}_1 \subset \mathfrak{b}_2 \subset \cdots$ es una sucesión estrictamente ascendente de ideales de B y $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$ es la sucesión también estrictamente ascendente de ideales de A dada por $\mathfrak{a}_i = f^{-1}(\mathfrak{b}_i)$ para toda i, entonces $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$ es finita, por lo que $\mathfrak{b}_1 \subset \mathfrak{b}_2 \subset \cdots$ también lo es. □

Teorema 2.70 (de la base de Hilbert) *Sea A un anillo Noetheriano. Entonces $A[X]$ también lo es.*

Demostración: Sea I un ideal de $A[X]$. Veremos que I es finitamente generado. Para cada $i \in \mathbb{N}$, sea

$$\mathfrak{a}_i = \{a \in A \mid \text{existe } f(X) = aX^i + \cdots + a_1X + a_0 \in I\}.$$

Es claro que cada $\mathfrak{a}_i$ es un ideal de A y que $\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$, por lo que existe $n \in \mathbb{N}$ tal que $\mathfrak{a}_n = \mathfrak{a}_{n+1} = \cdots$. Como cada ideal $\mathfrak{a}_i$ es finitamente generado, podemos escribir $\mathfrak{a}_i = (a_{i1}, ..., a_{im(i)})$, de manera que exista $f_{ij} \in I$ de grado i con coeficiente líder a_{ij}. Se afirma que $I = (f_{ij} \mid 1 \leq i \leq n,\ 1 \leq j \leq m(i))$. Veremos por inducción en $d = \operatorname{gr} f$, que $f \in I \Rightarrow f \in (f_{ij})$.

Si $d > n$, entonces existen $c_1, ..., c_{m(n)} \in A$ tales que el grado de f es mayor que el grado de $f - c_1 X^{d-n} f_{n1} - \cdots - c_{m(n)} X^{d-n} f_{nm(n)} \in I$. Si $d \leq n$, entonces existen $c'_1, ..., c'_{m(d)} \in A$ tales que el grado de f es mayor que el grado de $f - c'_1 f_{d1} - \cdots - c'_{m(d)} f_{dm(d)} \in I$. □

Un ideal $\mathfrak{a}$ es **irreducible** cuando $(\mathfrak{a} = \mathfrak{b} \cap \mathfrak{c}) \Rightarrow (\mathfrak{a} = \mathfrak{b}$ ó $\mathfrak{a} = \mathfrak{c})$. Un ideal $\mathfrak{q}$ de un anillo conmutativo R es **primario** cuando todo divisor de cero de $R/\mathfrak{q}$ es nilpotente. Supondremos que nuestros anillos son conmutativos.

Proposición 2.71 *Sea R un anillo Noetheriano. Entonces todo ideal se puede expresar como una intersección finita de ideales irreducibles.*

Demostración: Supongamos que $\Sigma \neq \varnothing$ es el conjunto de ideales que no admiten tal descomposición. Entonces Σ tiene un elemento máximo $\mathfrak{a}$, para el que existen ideales $\mathfrak{b}$ y $\mathfrak{c}$ tales que $\mathfrak{a} = \mathfrak{b} \cap \mathfrak{c}$ con $\mathfrak{a} \subset \mathfrak{b}$ y $\mathfrak{a} \subset \mathfrak{c}$. Pero entonces tanto $\mathfrak{b}$ como $\mathfrak{c}$ admiten tal descomposición, implicando que $\mathfrak{a}$ también se descompone así. □

Proposición 2.72 *Todo ideal irreducible de un anillo Noetheriano es primario.*

Demostración: Es suficiente ver que si el ideal (0) es irreducible, entonces es primario: Supongamos que $ab = 0$ con $b \neq 0$. Consideremos la cadena de ideales $\mathrm{an}(a) \subseteq \mathrm{an}(a^2) \subseteq \cdots$, que se estabiliza, digamos que $\mathrm{an}(a^n) = \mathrm{an}(a^{n+1})$.

Afirmamos que $(a^n) \cap (b) = (0)$: Esto es porque $x \in (b) \Rightarrow xa = 0$, por lo que si además $x \in (a^n)$, entonces $x = ya^n$, de manera que $ya^{n+1} = 0$, y así $y \in \mathrm{an}(a^{n+1}) = \mathrm{an}(a^n)$. Entonces $ya^n = 0$, teniéndose $x = 0$.

Como (0) es irreducible y $(b) \neq 0$, se obtiene $(a^n) = 0$, que demuestra que (0) es primario. □

Corolario 2.73 *Sea R un anillo Noetheriano. Entonces todo ideal se puede expresar como una intersección finita de ideales primarios.*

Ejercicios

1. Dé un ejemplo de un anillo Noetheriano que no sea principal.
2. Determine si el $\mathbb{Z}$-módulo $\mathbb{Q}$ es Noetheriano.
3. Demuestre que el anillo $M_n(R)$ de las matrices $n \times n$ sobre un anillo Noetheriano R es Noetheriano.
4. Dé un ejemplo de un anillo conmutativo que no sea Noetheriano.
5. Sean A un anillo Noetheriano y $f : A \to A$ un morfismo suprayectivo de anillos. Demuestre que f es inyectivo. (Sugerencia: Considere la cadena $\ker f \subseteq \ker f^2 \subseteq \cdots$).

2.10 Series Formales de Potencias

Sea A un anillo conmutativo con 1. El anillo de **series formales de potencias** $R = A[[X]]$ es el conjunto de las sucesiones $(a_0, a_1, a_2, ...)$, donde $a_i \in A$ para todo $i \in \mathbb{N}$, con operaciones suma y multiplicación dadas por

$$(a_0, ..., a_n, ...) + (b_0, ..., b_n, ...) = (a_0 + b_0, ..., a_n + b_n, ...).$$
$$(a_0, ..., a_n, ...) \times (b_0, ..., b_n, ...) = (c_0, ..., c_n, ...), \text{ con } c_n = \sum_{i+j=n} a_i b_j.$$

Se verifica inmediatamente que R es un anillo conmutativo; en el que $1 = (1, 0, 0, ...)$. Escribimos $X = (0, 1, 0, 0, ...)$, de manera que tenemos $X^2 = (0, 0, 1, 0, 0, ...)$, etc. Se dice que $\alpha = (a_0, a_1, a_2, ...)$ es de **orden** i cuando $a_0 = \cdots = a_{i-1} = 0,\ a_i \neq 0$, escrito $\circ(\alpha) = i$. No definimos el orden de 0.

Observaciones. Es inmediato que:

a) $\circ(\alpha + \beta) \geq \min\{\circ(\alpha), \circ(\beta)\}$, siempre que $\alpha, \beta, \alpha + \beta \neq 0$.

b) $\circ(\alpha\beta) \geq \circ(\alpha) + \circ(\beta)$, siempre que $\alpha, \beta, \alpha\beta \neq 0$.

c) Si A es un dominio, entonces $\circ(\alpha\beta) = \circ(\alpha) + \circ(\beta)$.

d) Si A es un dominio, entonces R también lo es.

Convenimos escribir $\sum_{i \geq 0} a_i X^i$ en lugar de $(a_0, a_1, a_2, ...)$, entendiendo que esta no es una suma, aunque $A[X]$ es un subanillo de $A[[X]]$.

Teorema 2.74 *Si A es un anillo Noetheriano, entonces el anillo de series formales $R = A[[X]]$ también es Noetheriano.*

Demostración: Sea I un ideal de R. Para cada $i \in \mathbb{N}$, definimos

$$\mathfrak{a}_i = \{a \in A \mid \text{existe } \alpha \in I, \alpha = aX^i + \text{ (términos de orden mayor)}\} \cup \{0\}.$$

Como A es Noetheriano, la cadena de ideales de A : $\mathfrak{a}_0 \subseteq \mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$ se estabiliza en algún punto, digamos que $\mathfrak{a}_n = \mathfrak{a}_{n+1}$, supongamos que $\mathfrak{a}_i = (a_{i1}, ..., a_{im(i)})$ para $1 \leq i \leq n$; y que

$$f_{ij} = a_{ij}X^i + (\text{términos de orden mayor}) \in I.$$

Afirmamos que $I = (f_{ij} \mid 1 \leq i \leq n,\ 1 \leq j \leq m(i))$. Si $d \leq n$ y $f = bX^d + (\text{términos de orden mayor}) \in I$ con $b \neq 0$, entonces existen $c_{d1}, ..., c_{dm(d)} \in A$ tales que $f - c_{d1}f_{d1} - \cdots - c_{dm(d)}f_{dm(d)} \in I$ es de orden mayor a d.

Si $d > n$, entonces también podemos encontrar $c_{d1}, ..., c_{dm(n)} \in A$ tales que $f - c_{d1}X^{d-n}f_{n1} - \cdots - c_{dm(n)}X^{d-n}f_{nm(n)} \in I$ es de orden mayor a d.

En todo caso, podemos suponer que $d > n$ para escribir

$$f = f_{n1}\sum_{k>n} c_{k1}X^{k-n} + \cdots + f_{nm(n)}\sum_{k>n} c_{km(n)}X^{k-n} \in (f_{ij}).$$

Como todo ideal de R es finitamente generado, se tiene que el anillo R es Noetheriano. □

Teorema 2.75 *Sean k un campo y $R = k[[X]]$. Entonces:*

a) $\sum_{i\geq 0} a_i X^i \in R^\star \Leftrightarrow a_0 \neq 0$.

b) Si $0 \neq \alpha \in R$, *entonces existe* $u \in R^\star$ *único tal que* $\alpha = uX^{\circ(\alpha)}$.

c) R es un dominio local con ideal máximo $\mathfrak{m} = (X)$ *y campo cociente* $R/\mathfrak{m} \cong k$.

d) R es un dominio principal.

e) El campo de fracciones de R, escrito $k((X))$, consiste de los elementos de la forma uX^j *con* $u \in R^\star$ *y* $j \in \mathbb{Z}$. *Además, sus operaciones son:*

$$\sum_{i\geq r} a_i X^i + \sum_{i\geq r} b_i X^i = \sum_{i\geq r} (a_i + b_i) X^i,$$

$$\sum_{i\geq r} a_i X^i \times \sum_{j\geq s} b_j X^j = \sum_{t\geq r+s} c_t X^t,$$

donde $c_t = \sum_{i+j=t} a_i b_j$ *y* $r, s \in \mathbb{Z}$.

Demostración: *a)* Si $\sum_{i\geq 0} a_i X^i \in R^\star$, entonces existe $\sum_{i\geq 0} b_i X^i$ tal que

$$(\sum_{i\geq 0} a_i X^i)(\sum_{i\geq 0} b_i X^i) = 1;$$

pero entonces $a_0 b_0 = 1 \Rightarrow a_0 \neq 0$.

Recíprocamente, supongamos que $\alpha = \sum_{i\geq 0} a_i X^i$ con $a_0 \neq 0$. Buscamos $\beta = \sum_{i\geq 0} b_i X^i$ tal que $\alpha\beta = 1$. Esta última igualdad es equivalente a la colección de ecuaciones:

$$\begin{array}{lllll}
a_0 b_0 & & & = & 1 \\
a_0 b_1 + a_1 b_0 & & & = & 0 \\
& \cdots & & = & \cdots \\
a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 & & & = & 0 \\
& \cdots & & = & \cdots
\end{array}$$

Este es un sistema de ecuaciones en las incógnitas b_i, que admite una solución única, pues b_0 se determina de la primera ecuación, b_1 de la segunda ecuación, etc. Todo esto gracias a que $a_0 \neq 0$.

b) Esta es una consecuencia inmediata de *a)*.

c) Esta también es una consecuencia inmediata de *a)*.

d) Si $0 \neq \alpha \in R$, entonces tenemos la igualdad de ideales $(\alpha) = (X^{\circ(\alpha)})$. Como R es Noetheriano, es suficiente observar que $(\alpha_1, ..., \alpha_r) = (X^s)$, donde $s = \min\{\circ(\alpha_i) \mid \alpha_i \neq 0\}$, suponiendo que algún $\alpha_i \neq 0$.

e) Si $\alpha = uX^{\circ(\alpha)}$ y $0 \neq \beta = vX^{\circ(\beta)}$ con $u, v \in R^\star$, entonces

$$\frac{\alpha}{\beta} = uv^{-1} X^{\circ(\alpha)-\circ(\beta)},$$

de donde se obtiene la primera afirmación de *e)*. Es fácil verificar que las operaciones de $k((X))$ son las enunciadas. □

Observaciones. Las siguientes afirmaciones son generalizaciones sencillas de lo anterior:

- $k[[X]]$ es un dominio de factorización única cuyo único irreducible es X junto con sus asociados:

$$(\sum_{i\geq 0} a_i X^i \text{ es irreducible}) \Leftrightarrow (a_0 = 0, a_1 \neq 0).$$

- Si A es un anillo conmutativo, entonces

$$\sum_{i\geq 0} a_i X^i \in A[[X]]^\star \Leftrightarrow a_0 \in A^\star.$$

Sean k un campo y $K = k(X_1, ..., X_n)$ el campo de funciones racionales sobre k en n variables. Para cada $i \in \mathbb{N} \setminus \{0\}$, definimos el polinomio simétrico $\rho_i = X_1^i + \cdots + X_n^i$.

Sean $\sigma_1, ..., \sigma_n$ los polinomios simétricos elementales. Consideremos al polinomio

$$f(T) = \prod_{i=1}^{n}(T - X_i) = T^n - \sigma_1 T^{n-1} + \cdots + (-1)^n \sigma_n \tag{2.6}$$

como elemento de $K((T^{-1}))$. Esto es posible, porque siendo un polinomio en T involucra solamente un número finito de potencias positivas de T.

Lema 2.76 *Sean k un campo y $a \in k$, entonces en $k((T^{-1}))$ tenemos que*

$$(T - a)^{-1} = T^{-1} + aT^{-2} + a^2T^{-3} + \cdots$$

Demostración: Es suficiente observar que

$$(T - a)(T^{-1} + aT^{-2} + a^2T^{-3} + \cdots) = 1$$

en $k((T^{-1}))$, según la multiplicación del Teorema 2.75 *e*). □

Teorema 2.77 (Newton) *Las siguientes identidades son válidas:*

$$\rho_r - \rho_{r-1}\sigma_1 + \cdots + (-1)^{r-1}\rho_1\sigma_{r-1} + (-1)^r r\sigma_r = 0, \text{ si } r \leq n. \tag{2.7}$$

$$\rho_r - \rho_{r-1}\sigma_1 + \cdots + (-1)^{n-1}\rho_{r-n+1}\sigma_{n-1} + (-1)^n\rho_{r-n}\sigma_n = 0, \text{ si } r > n. \tag{2.8}$$

Demostración: A partir de la ecuación (2.6), se tiene que

$$\frac{f'(T)}{f(T)} = \sum_{i=1}^{n}\frac{1}{T - X_i} = \sum_{i=1}^{n}(T^{-1} + X_iT^{-2} + \cdots) = nT^{-1} + \rho_1T^{-2} + \rho_2T^{-3} + \cdots$$

Para obtener la segunda igualdad usamos el lema precedente. Multiplicamos la identidad obtenida por

$$f(T)T^{1-n} = T(1 - \sigma_1 T^{-1} + \cdots + (-1)^n \sigma_n T^{-n}),$$

para obtener

$$f'(T)T^{1-n} = (1 - \sigma_1 T^{-1} + \cdots + (-1)^n \sigma_n T^{-n})(n + \rho_1 T^{-1} + \cdots). \quad (2.9)$$

Un cálculo directo, desde la ecuación (2.6), produce

$$f'(T)T^{1-n} = n - (n-1)\sigma_1 T^{-1} + \cdots + (-1)^{n-1}\sigma_{n-1}T^{-(n-1)}. \quad (2.10)$$

Comparando los coeficientes de las distintas potencias de T en (2.9) y (2.10), obtenemos las identidades (2.7) y (2.8) deseadas. □

Ejercicios

1. Sean k un campo y $R = k[[X]]$. Demuestre que R es un dominio Euclideano.

2. Sean k un campo, $A = k[[X]]$ y B el subconjunto de A formado por las series de forma $1 + \sum_{n\geq 1} a_n X^n$. Demuestre que dados un entero positivo r tal que $\text{caract}\, k \nmid r$ y $\alpha \in B$, existe un único $\beta \in B$ tal que $\beta^r = \alpha$.

3. Sea k un campo. Defina al anillo de series $R = k[[X_1, ..., X_n]]$; y demuestre que R es un dominio Noetheriano local.

2.11 Ejercicios Generales

1. Sean k un campo, $f(X) = a_0 X^m + a_1 X^{m-1} + \cdots + a_m$ un polinomio sobre k de grado menor o igual a m; y $g(X) = b_0 X^n + b_1 X^{n-1} + \cdots + b_n$ otro polinomio sobre k de grado menor o igual a n. Considere su resultante $R_{m,n}(f, g)$, dado por el determinante $(m+n) \times (m+n)$:

$$\begin{vmatrix} a_0 & \cdots & a_m & & \\ & \ddots & & \ddots & \\ & & a_0 & \cdots & a_m \\ b_0 & \cdots & b_n & & \\ & \ddots & & \ddots & \\ & & b_0 & \cdots & b_n \end{vmatrix}$$

Identificamos el espacio vectorial de los polinomios de grado menor o igual a m con k^{m+1} a través del isomorfismo

$$a_0X^m + a_1X^{m-1} + \cdots + a_m \longleftrightarrow (a_0, a_1, ..., a_m).$$

Demuestre que $R_{m,n} : k^{m+1} \times k^{n+1} \to k$ es la única función que satisface:

a) $R_{m,n}(f, g) = (-1)^{mn} R_{n,m}(g, f)$.

b) $R_{m,n}(\lambda f, \mu g) = \lambda^n \mu^m R_{m,n}(f, g)$ para todos $\lambda, \mu \in k$.

c) $R_{0,n}(a_0, g) = a_0^n$, independientemente de $g(X)$, para $m = 0$.

d) Cuando $m = 1$ y $n = 1$ se tiene que

$$R_{1,1}(aX + b, cX + d) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

e) Cuando $m = 1$ se tiene que $R_{1,n}(X - \rho, g) = g(\rho)$.

f) Se tiene bimultiplicatividad:

$$R_{m_1+m_2,n}(f_1 f_2, g) = R_{m_1,n}(f_1, g) R_{m_2,n}(f_2, g).$$
$$R_{m,n_1+n_2}(f, g_1 g_2) = R_{m,n_1}(f, g_1) R_{m,n_2}(f, g_2).$$

2. Sean k un campo y $f(X) = \prod_{i=1}^{n}(X - \alpha_i) \in k[X]$.

a) Evalúe el determinante de **Vandermonde**:

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} = \prod_{i>j}(\alpha_i - \alpha_j).$$

b) Escribiendo $b_j = \sum_{i=1}^{n} \alpha_i^j$ y $D = \operatorname{discr} f(X)$, demuestre que

$$\begin{vmatrix} n & b_1 & \cdots & b_{n-1} \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & \cdots & \vdots \\ b_{n-1} & b_n & \cdots & b_{2n-2} \end{vmatrix} = \prod_{i>j}(\alpha_i - \alpha_j)^2 = D.$$

c) Use el resultado anterior para demostrar que el discriminante de $X^3 - a_1X^2 + a_2X - a_3$ es $-4a_1^3a_3 + a_1^2a_2^2 + 18a_1a_2a_3 - 4a_2^3 - 27a_3^2$.

3. Sea R un anillo conmutativo y sea $Spec\ R$ el conjunto de los ideales primos de R. Para $C \subseteq R$, definimos $Z(C) = \{\mathfrak{p} \in Spec\ R \mid \mathfrak{p} \supseteq C\}$. Demuestre que

a) $Z(0) = Spec\ R$ y $Z(1) = \varnothing$.

b) $Z(\mathfrak{a} \cap \mathfrak{b}) = Z(\mathfrak{ab}) = Z(\mathfrak{a}) \cup Z(\mathfrak{b})$, para ideales $\mathfrak{a}$ y $\mathfrak{b}$.

c) $Z(\cup_i C_i) = \cap_i Z(C_i)$, para subconjuntos C_i.

Capítulo 3
Campos y Teoría de Galois

3.1 Extensiones de Campos

Para un campo k, existe un único morfismo de anillos $f : \mathbb{Z} \to k$ con $f(1) = 1$. Si $\ker f = (n)$, entonces la **característica** de k es n.

Como $\mathbb{Z}/n\mathbb{Z} \hookrightarrow k$, vemos que n es cero o bien es un número primo p. En el primer caso, k contiene una copia de $\mathbb{Z}$ y una copia de $\mathbb{Q}$. En el segundo caso, k contiene una copia de $\mathbb{Z}/p\mathbb{Z}$.

El **campo primo** de k es el subcampo generado por 1. Esto es, $\mathbb{Q}$ cuando $\operatorname{caract} k = 0$, ó bien $\mathbb{Z}/p\mathbb{Z}$ cuando $\operatorname{caract} k = p$.

Cuando k es un subcampo de K, se dice que K es una **extensión** de k; y se escribe K/k. En estas condiciones, K es un espacio vectorial sobre k. El **grado** de la extensión, escrito $[K : k]$, es la dimensión de este espacio vectorial. Se dice que la extensión es **finita** o **infinita**, según lo sea su grado.

Sea $f(X) \in k[X]$ irreducible y de grado positivo. Entonces, existe una raíz de $f(X)$ en k si y sólo si $\operatorname{gr} f = 1$. Para cualquier grado positivo de $f(X)$, se tiene que $k[X]/(f(X))$ es un campo K que contiene una copia de k y que está generado por la imagen α de X en el cociente. Esto se escribe así: $K = k(\alpha)$; y se dice que K se obtiene a partir de k adjuntándole α, raíz del polinomio $f(X)$.

Dada una extensión F/k, se dice que un elemento $\alpha \in F$ es **algebraico** sobre k cuando existen elementos $b_0, b_1, ..., b_n \in k$ no todos cero tales que

$$b_n\alpha^n + \cdots + b_1\alpha + b_0 = 0.$$

Esto equivale a decir que el morfismo de anillos $\psi : k[X] \to F$ tal que $\psi(X) = \alpha$ y que coincide con la identidad en k, tiene núcleo $I \neq (0)$. En esta situación, el núcleo tiene que ser de la forma $I = (p(X))$ con $p(X)$ irreducible. Si además $p(X)$ es mónico, decimos que $p(X)$ es el polinomio mínimo de α sobre k, escrito $\operatorname{Polmin}(\alpha, k)$. Cuando $I = (0)$, se dice que $\alpha \in F$ es **trascendente** sobre k; esto es equivalente a tener $k[X] \cong k[\alpha]$.

La extensión F/k es **algebraica** cuando todo elemento $\alpha \in F$ es algebraico sobre k. En caso contrario, la extensión es **trascendente**.

Dados una extensión F/k y un subconjunto $S \subseteq F$, el anillo generado por k y S se escribe $k[S]$, mientras que el campo generado por k y S se escribe $k(S)$. Cuando S consiste de un solo elemento, la situación es la siguiente:

Proposición 3.1 *Sean F/k una extensión de campos y $\alpha \in F$, entonces las siguientes condiciones son equivalentes:*
a) α es algebraico sobre k.
b) $k[\alpha] = k(\alpha)$.
c) $k(\alpha)/k$ es finita.
Cuando se satisfacen estas condiciones, se tiene la igualdad

$$[k(\alpha) : k] = \operatorname{gr} \operatorname{Polmin}(\alpha, k).$$

Demostración: *a)* $\Rightarrow$ *b)* : Si α es algebraico sobre k, entonces $k[\alpha] \cong k[X]/I$ con $(0) \neq I = (p(X))$ ideal primo de $k[X]$. Esto implica que I es máximo, que $k[\alpha]$ es un campo; y que $k[\alpha] = k(\alpha)$.

b) $\Rightarrow$ *c)* : Si $k[\alpha] = k(\alpha)$, entonces el anillo $k[\alpha]$ no es un anillo de polinomios, por lo que α es algebraico sobre k. Supongamos que n es el grado de $p(X) = \operatorname{Polmin}(\alpha, k)$. Afirmamos que $A = \{1, \alpha, \alpha^2, ..., \alpha^{n-1}\}$ es una base sobre k del espacio vectorial $k[\alpha] = k(\alpha)$:

Una relación de dependencia lineal de A es una ecuación polinomial de grado menor a n que α satisface, por lo que no existe.

Si $f(\alpha)$ es un polinomio en α con coeficientes en k, entonces por el algoritmo Euclideano, existen polinomios $q(X), r(X) \in k[X]$ tales que $f(X) = q(X)p(X) + r(X)$ con $r(X) = 0$ ó con $\operatorname{gr} r < n = \operatorname{gr} p$. Así, como $p(\alpha) = 0$, se tiene que $f(\alpha) = r(\alpha) \in$ espacio generado por A.

c) $\Rightarrow$ *a)* : Si $k(\alpha)/k$ es finita, entonces $\{1, \alpha, \alpha^2, ...\}$ es linealmente dependiente sobre k; y α es algebraico. □ La implicación *c)* $\Rightarrow$ *a)* tiene como consecuencia inmediata el siguiente resultado:

Corolario 3.2 *Toda extensión finita es algebraica.*

Observación. Pronto veremos que el recíproco de este corolario es falso.

Teorema 3.3 *Sean $k \subseteq F \subseteq K$ campos, $\{\alpha_i\}_{i \in I}$ una base de F/k y $\{\beta_j\}_{j \in J}$ una base de K/F, entonces $\{\alpha_i \beta_j\}_{(i,j) \in I \times J}$ es una base de K/k. En particular, $[K : k] = [K : F][F : k]$.*

Demostración: Esta última igualdad la entendemos así: La extensión K/k es finita si y sólo si K/F y F/k son finitas, en cuyo caso vale la igualdad escrita. Todo esto es consecuencia inmediata de la primera afirmación, que demostramos a continuación.

Dado $\gamma \in K$, existen $c_j \in F$ casi todos cero (= todos cero con un número finito de excepciones) tales que $\gamma = \sum c_j \beta_j$. Para cada c_j, existen $b_{ij} \in k$ casi todos cero tales que $c_j = \sum b_{ij} \alpha_i$. Así se tiene que $\gamma = \sum_{i,j} b_{ij} \alpha_i \beta_j$; y que $\{\alpha_i \beta_j\}$ genera a K sobre k.

Finalmente, si $\sum_{i,j} a_{ij}\alpha_i\beta_j = 0$, con $a_{ij} \in k$ casi todos cero, entonces para cada j se tiene que $\sum_{i,j} a_{ij}\alpha_i = 0$, pues las β_j son linealmente independientes sobre F. Esto implica que todo a_{ij} es cero; y que $\{\alpha_i\beta_j\}$ es una base de K/k. □

Los dos resultados siguientes proveen de ejemplos pertinentes.

Proposición 3.4 *Sean $\{p_1, ..., p_n, q_1, ..., q_m\}$ un conjunto con $n+m$ primos distintos y $K = \mathbb{Q}(\sqrt{p_1}, ..., \sqrt{p_n})$. Entonces $\sqrt{q_1 \cdots q_m} \notin K$.*

Demostración: Procedemos por inducción en n. Cuando $n = 0$, se tiene que $K = \mathbb{Q}$; y entonces $X^2 - q_1 \cdots q_m$ es irreducible en $\mathbb{Q}[X]$ por Eisenstein. Esto es, $\sqrt{q_1 \cdots q_m} \notin K$.

Si $F = \mathbb{Q}(\sqrt{p_1}, ..., \sqrt{p_{n-1}})$, entonces $\sqrt{p_n} \notin F$ por la hipótesis inductiva; y así $[K : F] = 2$. Si suponemos que $\sqrt{q_1 \cdots q_m} \in K$, entonces existen $a, b \in F$ con $\sqrt{q_1 \cdots q_m} = a + b\sqrt{p_n}$; y así $q_1 \cdots q_m = a^2 + 2ab\sqrt{p_n} + b^2 p_n$. Pero

$$ab \neq 0 \Rightarrow \sqrt{p_n} = \frac{q_1 \cdots q_m - a^2 - b^2 p_n}{2ab} \in F,$$

que es una contradicción. Los casos $b = 0$ y $a = 0$ implican respectivamente que $\sqrt{q_1 \cdots q_m} = a \in F$ ó que $\sqrt{q_1 \cdots q_m} = b\sqrt{p_n}$. En el último caso, se tiene $\sqrt{q_1 \cdots q_m p_n} = bp_n \in F$. Estas contradicciones terminan la demostración. □

De lo anterior, obtenemos inmediatamente

Corolario 3.5 *Si $\{p_1, ..., p_n\}$ es una lista de n primos distintos, entonces $[\mathbb{Q}(\sqrt{p_1}, ..., \sqrt{p_n}) : \mathbb{Q}] = 2^n$. La extensión $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \sqrt{11}, ...)/\mathbb{Q}$ es infinita.*

Teorema 3.6 *Sea $F = k(\alpha_1, ..., \alpha_n)$ con cada α_i algebraico sobre k. Entonces F/k es una extensión finita. Toda extensión generada por elementos algebraicos es algebraica.*

Demostración: $k \subseteq k(\alpha_1) \subseteq k(\alpha_1, \alpha_2) \subseteq \cdots \subseteq k(\alpha_1, ..., \alpha_n) = F$ es una cadena de extensiones finitas. El Teorema 3.3 garantiza que F/k es una extensión finita; mientras que el Corolario 3.2 garantiza que es algebraica.

Si E/k es una extensión generada por elementos algebraicos y $\alpha \in E$, entonces existen $\alpha_1, ..., \alpha_n \in E$ algebraicos tales que $\alpha \in k(\alpha_1, ..., \alpha_n)$. Concluimos que α es algebraico sobre k. □

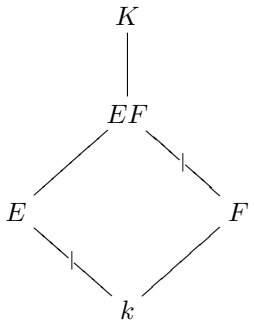
Corolario 3.7 *Si α y β son elementos algebraicos sobre el campo k, entonces también son algebraicos $\alpha \pm \beta$ y $\alpha\beta$; así como α/β cuando $\beta \neq 0$.*

Ejemplo. Sea $K = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}, ...)$. Aplicando los resultados previos, vemos que $K/\mathbb{Q}$ es una extensión algebraica infinita. Así, el recíproco del Corolario 3.2 es falso.

Una extensión F/k es **finitamente generada** cuando existen elementos $\alpha_1, ..., \alpha_n \in F$ tales que $F = k(\alpha_1, ..., \alpha_n)$. En el caso en que $n = 1$, la extensión es **simple** y α_1 es un **elemento primitivo**.

Si una extensión es finita, entonces es finitamente generada, porque una base es un conjunto de generadores. El recíproco es falso, pues si X es trascendente, entonces $k(X)/k$ es una extensión simple infinita.

Si E/k y F/k son extensiones para las que existe un campo K conteniendo tanto a E como a F, entonces el subcampo de K generado por $E \cup F$ se escribe EF. La extensión EF/F es la **translación** de E/k a F/k:



Sea $\mathcal{A}$ una familia de extesiones de campos. Diremos que $\mathcal{A}$ satisface la condición $\mathcal{T}$, por translación, cuando para E y F contenidos en un campo se tenga $E/k \in \mathcal{A} \Rightarrow EF/F \in \mathcal{A}$.

Similarmente, $\mathcal{A}$ satisface la condición $\mathcal{C}$, por cadena, cuando para toda cadena $k \subseteq E \subseteq F$ se tenga que $F/k \in \mathcal{A} \Leftrightarrow (F/E \in \mathcal{A}$ y $E/k \in \mathcal{A})$.

Si $\mathcal{A}$ es una familia de extesiones que satisface $\mathcal{T}$ y $\mathcal{C}$, entonces dadas dos extesiones $E/k, F/k \in \mathcal{A}$ tales que E y F estén contenidos en un campo, se tendrá $EF/k \in \mathcal{A}$.

Proposición 3.8 *Las extensiones finitas, así como las extensiones algebraicas satisfacen $\mathcal{T}$ y $\mathcal{C}$.*

Demostración: En el caso de las extensiones finitas, $\mathcal{C}$ se cumple por el Teorema 3.3.

Si E/k es una extensión finita y F/k es una extensión tal que E y F están contenidos en un campo, entonces existen $\alpha_1, ..., \alpha_n \in E$ tales que $E = k(\alpha_1, ..., \alpha_n)$ con todo α_i algebraico sobre k. Las cadenas de extensiones simples y finitas $k \subseteq k(\alpha_1) \subseteq k(\alpha_1, \alpha_2) \subseteq \cdots \subseteq k(\alpha_1, ..., \alpha_n) = E$ y $F \subseteq F(\alpha_1) \subseteq F(\alpha_1, \alpha_2) \subseteq \cdots \subseteq F(\alpha_1, ..., \alpha_n) = EF$ demuestran que es suficiente considerar el caso simple y finito: Aquí, $[F(\alpha) : F] \leq [k(\alpha) : k]$ porque α es raíz de $\text{Polmin}(\alpha, k) \in F[X]$.

En cuanto a las extensiones algebraicas, es obvio que $\mathcal{T}$ se cumple. Dada una cadena $k \subseteq E \subseteq F$, la implicación F/k algebraica $\Rightarrow$ (F/E algebraica y E/k algebraica) es clara. Recíprocamente, sean F/E, E/k extensiones algebraicas y $\alpha \in F$, entonces existen $b_m, ..., b_1, b_0 \in E$ algebraicos sobre k

no todos cero tales que $b_m\alpha^m + \cdots + b_1\alpha + b_0 = 0$, por lo que α es algebraico sobre $k(b_0, b_1, ..., b_m)$. Así, $[k(b_0, b_1, ..., b_m, \alpha) : k]$ es finito, por lo que α es algebraico sobre k y $\mathcal{C}$ se cumple. □

Como consecuencia inmediata tenemos:

Corolario 3.9 *Si E/k es una extensión finita y F/k es una extensión tal que E y F están contenidos en un campo, entonces $[EF : F] \leq [E : k]$.*

Observaciones.

1. Para la familia de las extensiones finitamente generadas, es claro que $\mathcal{T}$ se satisface. Aunque la condición $\mathcal{C}$ también se cumple; este resultado no lo demostramos ni lo utilizamos.

2. Si F y K son campos y $\varphi : F \to K$ es un morfismo (de anillos), entonces $\varphi(1) = 1$ y $\ker\varphi$ es un ideal de F, por tanto $\ker\varphi = (0)$. Así, todos los morfismos de campos son inyectivos.

3. El estudio de morfismos de campos se reduce al estudio de inclusiones de un campo dentro de otro campo.

Si F y K son extensiones de un campo k y $\varphi : F \to K$ es un morfismo tal que restringido a k es la identidad, decimos que φ es un k-**morfismo**.

Proposición 3.10 *Sean F/k extensión algebraica y $\varphi : F \to F$ un k-morfismo, entonces φ es biyectivo.*

Demostración: Ya sabemos que φ es inyectivo. Sean $\alpha \in F$ arbitrario, $p(X) = \text{Polmin}(\alpha, k)$ y sea E el subcampo de F generado por las raíces de $p(X)$. Es suficiente ver que $\alpha \in \text{Im}\,\varphi$.

La extensión E/k es finita y $\varphi(E) \subseteq E$. Como $\varphi|_E$ es una transformación k-lineal e inyectiva del espacio vectorial E sobre sí mismo, se tiene que $\varphi|_E$ es suprayectivo; y que $\alpha \in \text{Im}\,\varphi$. □

Ejercicios

1. Demuestre que los campos $\mathbb{Q}(\sqrt{2})$ y $\mathbb{Q}(\sqrt[3]{3})$ no son isomorfos.

2. Sean $a, b \in K$ algebraicos sobre k, de grados m y n respectivamente, con $(m, n) = 1$. Demuestre que $[k(a, b) : k] = mn$.

3. Encuentre un elemento $\beta \in \mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$ tal que $\mathbb{Q}(\beta) = \mathbb{Q}(\sqrt{2}, \sqrt[3]{3})$.

4. Sean F/k una extensión algebraica y D un dominio con $k \subseteq D \subseteq F$. Demuestre que D es un campo.

5. Sea K/k una extensión finita de campos tal que si E y F son campos intermedios, se tiene $E \subseteq F$ ó bien $F \subseteq E$. Demuestre que K/k es simple.

6. Sea $k(\alpha)/k$ una extensión de grado 5. Demuestre que $k(\alpha) = k(\alpha^3)$.

7. Sea k un campo de característica cero. Demuestre que el número de polinomios mónicos irreducibles de grado 2 en $k[X]$ es cero o infinito.

8. Sean $f(X)$ un polinomio irreducible sobre un campo k y $a \in k$. Demuestre que $f(X)$ permanece irreducible sobre $k(\sqrt{a})$ ó bien se descompone como el producto de dos factores del mismo grado.

3.2 Cerradura Algebraica

Proposición 3.11 *Las siguientes condiciones en un campo k son equivalentes:*
a) Si F/k es una extensión algebraica, entonces $F = k$.
b) Si F/k es una extensión finita, entonces $F = k$.
c) Todo polinomio irreducible en $k[X]$ es de grado uno.
d) Todo polinomio de $k[X]$ de grado positivo es un producto de polinomios lineales en $k[X]$.
e) Todo polinomio de $k[X]$ de grado positivo tiene una raíz en k.

La demostración de este resultado se deja como ejercicio. Se dice que un campo k es **algebraicamente cerrado** cuando satisface las condiciones de la proposición.

Proposición 3.12 *Todo campo algebraicamente cerrado es infinito.*

Demostración: Si k es un campo cuyos elementos son $a_1, ..., a_n$, entonces $(X - a_1) \cdots (X - a_n) + 1$ no tiene raíces en k. □

Pronto veremos que $\mathbb{C}$ es algebraicamente cerrado. A continuación nos proponemos construir campos algebraicamente cerrados a partir de un campo dado.

Teorema 3.13 *Dado un campo k, existe una extensión F/k con F algebraicamente cerrado.*

Demostración: Primero construiremos una extensión F_1/k tal que todo polinomio en $k[X]$ de grado positivo tenga una raíz en F_1.

A cada $f(X) \in k[X]$ de grado positivo le asociamos una variable X_f; y llamamos S al conjunto de dichas variables.

Afirmamos que en el anillo de polinomios $k[S]$, el conjunto $\{f(X_f)\}$ genera un ideal $\mathfrak{a}$, es decir, que $\mathfrak{a} \neq k[S]$. De no ser así, existirían elementos $g_1, ..., g_n \in k[S]$ tales que $g_1 f_1 + \cdots + g_n f_n = 1$, donde cada $f_i = f_i(X_i)$ es uno de nuestros generadores, cambiando ligeramente la notación. Ahora bien, cada g_i involucra un número finito de variables, así existe N tal que

$$1 = \sum_i g_i(X_1, ..., X_N) f_i(X_i). \tag{3.1}$$

Esto es absurdo: Si K es una extensión de k conteniendo α_i raíz de f_i para cada i, procedemos a evaluar (3.1) en K para llegar a $1 = 0$.

Dado que $\mathfrak{a} \neq k[S]$, existe un ideal máximo $\mathfrak{m}$ de $k[S]$ tal que $\mathfrak{a} \subseteq \mathfrak{m}$. Escribimos $F_1 = k[S]/\mathfrak{m}$, para tener que F_1/k es una extensión de k tal que todo polinomio $f(X_f) \in k[X]$ de grado positivo tiene una raíz en F_1: la imagen de X_f en el cociente.

Continuamos de la misma manera con F_1, etc. para obtener una cadena de campos $F_1 \subseteq F_2 \subseteq \cdots$ Definimos $F = \cup_i F_i$. Es fácil ver que F es un campo algebraicamente cerrado. □

Se dice que K es una **cerradura algebraica** de k cuando K/k es una extensión algebraica tal que K es algebraicamente cerrado.

Corolario 3.14 *Todo campo k, admite una cerradura algebraica.*

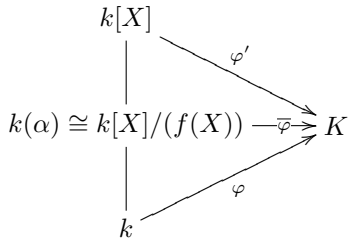
Demostración: Sabemos que existe F/k con F algebraicamente cerrado. Definimos $K = \{\alpha \in F \mid \alpha \text{ es algebraico sobre } k\}$. Entonces K es un campo por el Corolario 3.7, la extensión K/k es algebraica por el Teorema 3.6; y K es algebraicamente cerrado también por el Teorema 3.6. □

A continuación iniciamos un estudio del siguiente problema: Dados un morfismo de campos $\varphi : k \to K$ con K algebraicamente cerrado y una extensión algebraica F/k, ¿es posible extender φ a F? ¿cuántas extensiones existen?

Proposición 3.15 *Sean $\varphi : k \to K$ un morfismo de campos con K algebraicamente cerrado, $F = k(\alpha)$ una extensión simple algebraica y $f(X) = \text{Polmin}(\alpha, k)$. Entonces las extensiones de φ a F corresponden a las distintas raíces de $\varphi f(X)$ en K. Siempre existe al menos una extensión; y hay cuando más* $\text{gr}\, f(X)$ *extensiones.*

Demostración: En primer lugar, $f(\alpha) = 0 \Rightarrow \overline{\varphi} f(\alpha) = 0$ para cualquier extensión $\overline{\varphi}$ de φ, por lo que α tiene que ir a alguna raíz de $\varphi f(X)$ en K.

Recíprocamente, el morfismo φ puede extenderse a un morfismo de anillos $\varphi' : k[X] \to K$ al escribir $\varphi'(X) = \beta$ para cualquier $\beta \in K$.



Si β es raíz del polinomio $\varphi f(X) \in K[X]$, entonces φ' se factoriza a través de $\overline{\varphi}$ como en el diagrama. Las demás conclusiones son claras. □

Teorema 3.16 *Sean $\varphi : k \to K$ un morfismo de campos con K algebraicamente cerrado y F/k una extensión algebraica. Entonces φ se puede extender a F. Si además, F y K son cerraduras algebraicas de k, entonces F y K son isomorfos ante cualquiera de estas extensiones.*

Demostración: Sea $\mathcal{C}$ la clase de las parejas (E, ψ) tales que E es un campo con $k \subseteq E \subseteq F$ y $\psi : E \to K$ es una extensión de φ. Definimos un orden parcial en $\mathcal{C}$ así: $(E, \psi) \leq (E', \psi')$ cuando $E \subseteq E'$ y $\psi'|_E = \psi$.

Por el Lema de Zorn, existe una pareja máxima (E, ψ). Si $E \neq F$, entonces existe $\alpha \in F$ con $\alpha \notin E$, de manera que ψ se puede extender a $E(\alpha)$. Esta contradicción demuestra que $E = F$.

Cuando además F y K son cerraduras algebraicas de k, se tiene que K es algebraico sobre el campo algebraicamente cerrado $\varphi(F)$. Se concluye que $K = \varphi(F)$. □

Escribimos $\overline{k}$ para expresar la cerradura algebraica del campo k. El **campo de los números algebraicos** es $\overline{\mathbb{Q}}$.

Ejercicio

1. Demuestre que $\overline{\mathbb{Q}}$ es numerable, que $\overline{\mathbb{Q}} \neq \mathbb{C}$; y que $\overline{\mathbb{Q}}/\mathbb{Q}$ es una extensión algebraica infinita.

3.3 Normalidad

Sean k un campo y $f(X) \in k[X]$ de grado positivo. Se dice que F es un **campo de descomposición de $f(X)$ sobre k** cuando existe una factorización $f(X) = c\prod_{i=1}^{n}(X - \alpha_i)$ en $F[X]$ y además $F = k(\alpha_1, ..., \alpha_n)$.

Teorema 3.17 *Si K y F son dos campos de descomposición de $f(X)$ sobre k, entonces existe un k-isomorfismo $\varphi : F \to K$.*

Demostración: La inclusión $k \hookrightarrow \overline{K}$ admite una extensión al morfismo $\varphi : F \to \overline{K}$. Es suficiente ver que $\operatorname{Im}\varphi = K$.

Podemos suponer que $f(X)$ es mónico, que $K = k(\alpha_1, ..., \alpha_n)$, que $F = k(\beta_1, ..., \beta_n)$, que $f(X) = \prod_i (X - \alpha_i)$ en $K[X]$; y que $f(X) = \prod_i (X - \beta_i)$ en $F[X]$.

El resultado de aplicar φ a $f(X)$ es

$$\prod_{i=1}^{n}[X - \varphi(\beta_i)] = \varphi f(X) = f(X) = \prod_{i=1}^{n}(X - \alpha_i).$$

La factorización única del anillo $\overline{K}[X]$ implica la igualdad de conjuntos $\{\alpha_1, ..., \alpha_n\} = \{\varphi(\beta_1), ..., \varphi(\beta_n)\}$, por lo que $\varphi(F) = K$. □

Si $\{f_i(X)\}_{i\in I} \subseteq k[X]$ es una colección de polinomios de grado positivo, se dice que F es un **campo de descomposición de** esta colección sobre k cuando existe una factorización de cada $f_i(X)$ como producto de polinomios lineales en $F[X]$; y además F está generado sobre k por las raíces de los $f_i(X)$.

Corolario 3.18 *Si K y F son campos de descomposición sobre k de una misma colección de polinomios, entonces hay un k-isomorfismo $\varphi : F \to K$.*

Demostración: Para cada i hay un único campo de descomposición F_i de f_i en F y otro K_i en K.

Como F/k es una extensión algebraica, existe un k-morfismo $\varphi : F \to \overline{K}$ que satisface $\varphi(F_i) = K_i$ para toda i. Observando que F está generado por $\cup_i F_i$, mientras que K lo está por $\cup_i K_i$, tenemos que $\varphi(F) = K$. □

Teorema 3.19 *Para una extensión algebraica F/k, las siguientes condiciones son equivalentes:*

a) Todo elemento irreducible de $k[X]$ con una raíz en F, se factoriza como producto de polinomios lineales en $F[X]$.

b) F es el campo de descomposición de una colección de polinomios de $k[X]$.

c) Si suponemos que $F \subseteq \overline{k}$, entonces todo k-morfismo $\varphi : F \to \overline{k}$, se restringe a un automorfismo de F.

Demostración: *a)* $\Rightarrow$ *b)*: F es el campo de descomposición de la colección de los polinomios mínimos sobre k de todos los elementos de F.

b) $\Rightarrow$ *c)*: Si F es el campo de descomposición de $\{f_i(X)\}_{i\in I}$ sobre k, entonces F contiene un único subcampo de descomposición F_i de $f_i(X)$; y F está generado por $\cup_{i\in I} F_i$. Si φ es un k-morfismo de F en $\overline{k}$, entonces $\varphi(F_i) = F_i$ para toda i; y así $\varphi(F) = F$.

c) $\Rightarrow$ *a)*: Sea $g(X) \in k[X]$ irreducible, con una raíz $\alpha \in F$. Para otra raíz $\beta \in \overline{k}$ de $g(X)$, existe un k-morfismo $\varphi : k(\alpha) \to \overline{k}$, tal que $\varphi(\alpha) = \beta$. Este morfismo admite una extensión a F. La hipótesis implica que $\beta \in F$. □

Una extensión algebraica F/k es **normal** cuando satisface las condiciones del teorema anterior.

Ejemplo. Sean $X^2 - 2, X^4 - 2 \in \mathbb{Q}[X]$, α una raíz de $X^2 - 2$ y β una raíz de $X^4 - 2$ con $\beta^2 = \alpha$.

Los polinomios son irreducibles en $\mathbb{Q}[X]$ por el criterio de Eisenstein. Escribimos $K = \mathbb{Q}(\alpha)$ y $F = \mathbb{Q}(\beta)$, para tener $[K : \mathbb{Q}] = 2$, $[F : \mathbb{Q}] = 4$; y por tanto, $[F : K] = 2$.

$$\begin{array}{c} F = \mathbb{Q}(\beta) \\ \Big| \, 2 \\ K = \mathbb{Q}(\alpha) \\ \Big| \, 2 \\ \mathbb{Q} \end{array}$$

K es un campo de descomposición de $X^2 - 2$ sobre $\mathbb{Q}$, por lo que $K/\mathbb{Q}$ es una extensión normal. Análogamente, F/K es normal porque F es campo de descomposición de $X^2 - \alpha$ sobre K. Sin embargo, la extensión $F/\mathbb{Q}$ no es normal, como veremos a continuación.

Supongamos que $F/\mathbb{Q}$ es normal. El polinomio $f(X) = X^4 - 2$ tiene una raíz en F, por tanto se factoriza totalmente en $F[X]$. Como $f'(X) = 4X^3$ no tiene raíces comunes con $f(X)$, concluimos que $X^4 - 2$ tiene cuatro raíces distintas $\beta = \beta_1, \beta_2, \beta_3, \beta_4$ en F.

Sea $\zeta_i = \beta_i/\beta_1$ para $1 \leq i \leq 4$. Es inmediato que $\zeta_i^4 = 1$ para cada i. Por esto, $\zeta_1, \zeta_2, \zeta_3, \zeta_4$ son cuatro raíces distintas de $X^4 - 1 = (X^2 + 1)(X^2 - 1)$.

Ahora afirmamos que $X^2 + 1$ es irreducible en $F[X]$. Esta contradicción demostrará que $F/\mathbb{Q}$ no es normal. En efecto, si $c_1 + c_2\beta$ es raíz de $X^2 + 1$ con $c_1, c_2 \in K$, entonces $(c_1 + c_2\beta)^2 = (c_1^2 + c_2^2\beta^2) + 2c_1c_2\beta = -1$, y de ahí que $c_1^2 + c_2^2\alpha = -1$ y $2c_1c_2 = 0$. Por tanto, $c_1 = 0$ ó bien $c_2 = 0$.

Por un lado, $c_1 = 0 \Rightarrow c_2^2\alpha = -1$, mientras que $c_2 = 0 \Rightarrow c_1^2 = -1$. Veamos que ambas conclusiones son imposibles.

Escribiendo $c_1 = a_1 + a_2\alpha$ con $a_1, a_2 \in \mathbb{Q}$, la ecuación $c_1^2 = -1$ se transforma en $(a_1^2 + 2a_2^2) + 2a_1a_2\alpha = -1$, por lo que $(a_1^2 + 2a_2^2) = -1$, que no tiene soluciones en $\mathbb{Q}$.

Escribiendo $c_2 = b_1 + b_2\alpha$ con $b_1, b_2 \in \mathbb{Q}$, obtenemos de $c_2^2\alpha = -1$ la ecuación $4b_1b_2 + (b_1^2 + 2b_2^2)\alpha = -1$, y de ahí, $b_1^2 + 2b_2^2 = 0$ y $4b_1b_2 = -1$, que no admiten soluciones en $\mathbb{Q}$.

Este ejemplo demuestra que la familia de extensiones normales no satisface $\mathcal{C}$, los siguientes resultados nos dan propiedades de esta familia.

Proposición 3.20 *La familia de las extensiones normales satisface $\mathcal{T}$. Si $k \subseteq K \subseteq F$ con F/k normal, entonces F/K es normal. Si E_1 y E_2 son extensiones normales de k contenidas en algún campo, entonces E_1E_2/k y $(E_1 \cap E_2)/k$ también son normales.*

Demostración: Supongamos que F_1 y F_2 son extensiones de k contenidas en un campo; y que F_1/k es normal. Entonces F_1 es campo de descomposición sobre k de una colección de polinomios; y F_1F_2 también es campo de descomposición sobre F_2 de la misma colección de polinomios. Así, F_1F_2/F_2 es normal y $\mathcal{T}$ se cumple.

Si ahora F/k es normal y $k \subseteq K \subseteq F$, entonces F es campo de descomposición sobre k y también sobre K de una colección de polinomios en $k[X] \subseteq K[X]$.

Si E_1 y E_2 son extensiones normales de k contenidas en un campo, entonces E_1 es campo de descomposición de $\{f_i\} \subseteq k[X]$, E_2 lo es de $\{g_j\} \subseteq k[X]$, por lo que E_1E_2 es campo de descomposición de $\{f_i, g_j\}$. Así, E_1E_2/k es normal. Si $h(X) \in k[X]$ es irreducible y tiene una raíz en $E_1 \cap E_2$, entonces se factoriza totalmente en E_1 y en E_2, y por ello, en $E_1 \cap E_2$. Se concluye que $(E_1 \cap E_2)/k$ es normal. □

Proposición 3.21 *Sean F/k una extensión normal y $f(X)$ un elemento irreducible de $k[X]$ con factores mónicos irreducibles $g(X), h(X) \in F[X]$. Entonces existe un k-automorfismo φ de F tal que $\varphi(g(X)) = h(X)$.*

Demostración: Podemos suponer que $F \subseteq \overline{k}$. Sean α una raíz de $g(X)$ y β una raíz de $h(X)$. Entonces $g(X) = \text{Polmin}(\alpha, F)$, $h(X) = \text{Polmin}(\beta, F)$.

Como α y β son raíces de $f(X)$ que es irreducible sobre k, existe un k-morfismo $\varphi : k(\alpha) \to k(\beta)$ tal que $\varphi(\alpha) = \beta$.

Este morfismo puede extenderse a $\varphi : F \to \overline{k}$, que a su vez admite una restricción a un automorfismo de F por ser F/k normal; y que al actuar en los coeficientes de $g(X)$, se tiene que $\varphi(g(X)) = h(X)$, pues $\varphi(\alpha) = \beta$. □

Dada una extensión algebraica F/k con $F \subseteq \overline{k}$, definimos la **cerradura normal** de F/k en $\overline{k}$, como la extensión K de F que satisface las siguientes condiciones claramente equivalentes:

a) K es la intersección de todas las extensiones $E \subseteq \overline{k}$ de F con E/k normal.

b) K es el campo generado por $\cup_\sigma \sigma(F)$, donde σ varía sobre la colección de todos los k-morfismos de F en $\overline{k}$.

c) K es el campo de descomposición de los polinomios mínimos de un conjunto de generadores de F/k.

Observemos que si F/k es finita, entonces la cerradura normal K/k también lo es, gracias a la condición *c*).

Ejercicios

1. Construya un campo de descomposición K para $X^5 - 2$ sobre $\mathbb{Q}$. Calcule $[K : \mathbb{Q}]$.

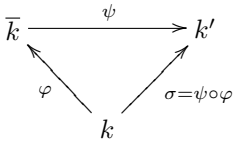
2. Construya un campo de descomposición F para $X^6 - 1$ sobre $\mathbb{Q}$. Calcule $[F : \mathbb{Q}]$.

3. Construya una extensión normal $K/\mathbb{Q}$ con $[K : \mathbb{Q}] = 3$.

4. Sea α una raíz de $13X^4 - 29X^2 + 13 \in \mathbb{Q}[X]$. Demuestre que $\mathbb{Q}(\alpha)/\mathbb{Q}$ es normal.

3.4 Separabilidad

Dada una extensión finita de campos F/k, definimos el **grado de separabilidad de** F/k, escrito $[F : k]_s$, como el número de k-morfismos $\varphi : F \to \overline{k}$, donde $\overline{k}$ es una cerradura algebraica de k. Este número no depende de la cerradura algebraica elegida, pues si k' fuera otra, entonces existiría un k-isomorfismo $\psi : \overline{k} \to k'$ que nos permitiría construir una biyección $\Psi : \mathrm{Hom}_k(F, \overline{k}) \to \mathrm{Hom}_k(F, k')$ del conjunto $\mathrm{Hom}_k(F, \overline{k})$ de k-morfismos de F en $\overline{k}$ al conjunto $\mathrm{Hom}_k(F, k')$ de k-morfismos de F en k' así: $\Psi(\varphi) = \sigma = \psi \circ \varphi$.



Proposición 3.22 *Sea F/k una extensión finita de campos. Entonces:*
a) $[F : k]_s \geq 1$.
b) Si $F = k(\alpha)$, entonces $[F : k]_s \leq \mathrm{gr}\,\mathrm{Polmin}(\alpha, k)$. *Además,* $[F : k]_s$ *es el número de raíces distintas de* $\mathrm{Polmin}(\alpha, k)$ *en* $\overline{k}$.
c) $[F : k]_s = [F : K]_s[K : k]_s$ *para cualquier campo intermedio* K.
d) $[F : k]_s \leq [F : k]$.

Demostración: *a)* y *b)* son reenunciados de la Proposición 3.15.

c) resulta de observar que todo k-morfismo $\varphi : F \to \overline{k}$ admite una restricción a K.

Sabemos que *d)* vale para el caso de extensiones simples. Como F/k es una extensión finita, existen $\alpha_1, ..., \alpha_n$ tales que $F = k(\alpha_1, ..., \alpha_n)$. La conclusión se obtiene de observar que $[k(\alpha_1, ..., \alpha_{i+1}) : k(\alpha_1, ..., \alpha_i)]_s \leq [k(\alpha_1, ..., \alpha_{i+1}) : k(\alpha_1, ..., \alpha_i)]$; y que ambos grados son multiplicativos. □

La extensión finita F/k se dice **separable** cuando $[F : k]_s = [F : k]$. Dados una extensión arbitraria de campos K/k y un elemento $\alpha \in K$ algebraico sobre k, decimos que α es **separable sobre** k cuando la extensión (finita) $k(\alpha)/k$ lo es.

Un polinomio en $k[X]$ se llama **separable** cuando no tiene raíces múltiples en $\overline{k}$. Por tanto, un elemento α, algebraico sobre k, es separable sobre k si y sólo si su polinomio mínimo lo es.

Proposición 3.23 *Sea F/k una extensión finita. Entonces F/k es separable si y sólo si todo elemento de F es separable sobre k.*

Demostración: Si F/k es separable y $\alpha \in F$, consideramos la cadena $k \subseteq k(\alpha) \subseteq F$. La multiplicatividad de los grados implica que $[k(\alpha) : k]_s = [k(\alpha) : k]$, esto es, que α es separable sobre k.

Recíprocamente, existen $\alpha_1, ..., \alpha_n \in F$ tales que $F = k(\alpha_1, ..., \alpha_n)$, con cada α_i separable sobre k. Por esto, todo Polmin(α_i, k) y su factor Polmin$(\alpha_i, k(\alpha_1, ..., \alpha_{i-1}))$ son separables. Así,

$$[k(\alpha_1, ..., \alpha_i) : k(\alpha_1, ..., \alpha_{i-1})]_s = [k(\alpha_1, ..., \alpha_i) : k(\alpha_1, ..., \alpha_{i-1})] \text{ para toda } i,$$

obteniéndose $[F : k]_s = [F : k]$ de la multiplicatividad de los grados. □

Este resultado nos permite definir "separabilidad" con mayor generalidad: Una extensión algebraica (no necesariamente finita) K/k es separable si y sólo si todo elemento de K es separable sobre k. La nueva definición generaliza a la anterior. Esta es la mayor generalidad para la que definimos "separabilidad".

Proposición 3.24 *Sea k un campo de característica cero. Todo polinomio irreducible en $k[X]$ es separable. En particular, toda extensión algebraica de k es separable.*

Demostración: Sean $0 \neq f(X) \in k[X]$ irreducible de grado positivo y $\alpha \in \overline{k}$ raíz de $f(X)$. Sabemos que (α es raíz múltiple) $\Leftrightarrow f'(\alpha) = 0$.

Como $f'(X)$ es de grado menor que el de $f(X)$, tenemos que (α raíz múltiple) $\Rightarrow f'(\alpha) = 0 \Rightarrow f'(X) = 0$, lo cual es absurdo porque la característica es cero. □

El resultado anterior dice claramente que los problemas de falta de separabilidad se dan solamente en característica positiva. Esta es la razón por la que muchos autores consideran solamente campos de característica cero. Por el contrario, supondremos hasta nuevo aviso, que la característica de los campos que se mencionen es $p > 0$.

Proposición 3.25 *Si el polinomio irreducible $f(X) \in k[X]$ no es separable, entonces $f(X)$ es un polinomio en X^p (con coeficientes en k).*

Demostración: En la demostración anterior vimos que la hipótesis implica que $f'(X) = 0$. Como $(aX^n)' = naX^{n-1} = 0$ con $a \neq 0$ es posible solamente con $p \mid n$, se obtiene la conclusión. □

Sea k un campo de característica p. Recordemos que

$$p \mid \binom{p}{i} \text{ para } 1 \leq i \leq p-1.$$

De manera que $(a+b)^p = a^p + b^p$ para todos $a, b \in k$. Como $(ab)^p = a^p b^p$ también vale para todos $a, b \in k$, tenemos que la función $\sigma : k \to k$ dada por $\sigma(a) = a^p$ para todo $a \in k$, es un morfismo de anillos, el llamado **morfismo de Frobenius**. Escribimos $k^p = \operatorname{Im} \sigma$.

Decimos que una extensión algebraica F/k, no necesariamente finita, es **inseparable pura** cuando $[F : k]_s = 1$, es decir, cuando para toda

cerradura algebraica $\overline{k}$ de k, existe un único k-morfismo $\varphi : F \to \overline{k}$. Un elemento $\alpha \in F$ es **inseparable puro sobre** k cuando $[k(\alpha) : k]_s = 1$. Es claro que una extensión algebraica F/k es inseparable pura si y sólo si todo elemento $\alpha \in F$ es inseparable puro sobre k.

Proposición 3.26 *Dada F/k una extensión algebraica, escribimos $A = \{\alpha \in F \mid \alpha$ separable sobre $k\}$ y $B = \{\beta \in F \mid \beta$ inseparable puro sobre $k\}$. Entonces A y B son campos y $A \cap B = k$.*

Demostración: Sean $a, b \in A$ con $b \neq 0$. Entonces b es separable sobre $k(a)$; y en la cadena $k \subseteq k(a) \subseteq k(a, b)$ se tiene que $[k(a) : k]_s = [k(a) : k]$ y que $[k(a, b) : k(a)]_s = [k(a, b) : k(a)]$. Así, $k(a, b)/k$ es separable. En particular, $a \pm b, ab, a/b$ son todos separables sobre k; y A es un campo.

De manera análoga se ve que B es un campo.

Finalmente, $c \in A \cap B \Rightarrow [k(c) : k] = [k(c) : k]_s = 1 \Rightarrow c \in k$. □

En las condiciones de la proposición anterior, decimos que A es la **cerradura separable** de k en F; y que B es la **cerradura inseparable pura** de k en F. Es claro que todo elemento de F separable sobre A está en A; y que todo elemento de F inseparable puro sobre B está en B.

Proposición 3.27 *Sea α un elemento algebraico sobre k. Las siguientes condiciones en α son equivalentes:*

a) α inseparable puro sobre k.

b) El polinomio mínimo de α sobre k es de la forma $X^{p^r} - b$.

c) α es raíz de un polinomio de la forma $X^{p^t} - c \in k[X]$.

Demostración: La implicación $b) \Rightarrow c)$ es clara.

Veamos que $c) \Rightarrow a)$: Si α es raíz de $X^{p^t} - c$, entonces $\alpha^{p^t} = c$ y así $(X - \alpha)^{p^t} = X^{p^t} - \alpha^{p^t} = X^{p^t} - c$, por lo que α es raíz de un polinomio en $k[X]$ con una sola raíz. Así, $[k(\alpha) : k]_s = 1$.

Para ver $a) \Rightarrow b)$, digamos que $f(X) = \text{Polmin}(\alpha, k)$. Si $\operatorname{gr} f(X) \geq 1$, entonces $f(X)$ no es separable y la Proposición 3.25 garantiza que existe $g(Y) \in k[Y]$ tal que $f(X) = g(X^p)$. Claramente, $g(Y)$ es irreducible y podemos repetir el proceso si $g(Y)$ no es separable. En algún momento llegamos a obtener un polinomio $h(Y) \in k[Y]$ separable e irreducible tal que $f(X) = h(X^{p^r})$. Sea $\beta = \alpha^{p^r}$, de manera que β es separable sobre k al ser raíz de $h(Y)$. De la cadena $k \subseteq k(\beta) \subseteq k(\alpha)$ y de $[k(\alpha) : k]_s = 1$, se obtiene que $\beta \in k$, es decir, que $h(Y)$ es lineal; y que $f(X) = X^{p^r} - b$. □

Teorema 3.28 *Las extensiones algebraicas separables y las inseparables puras satisfacen $\mathcal{T}$ y $\mathcal{C}$.*

Demostración: Supongamos que $k \subseteq F \subseteq K$ y que K/k es una extensión algebraica separable, entonces todo elemento de K es separable sobre F y

todo elemento de F es separable sobre k, de manera que K/F y F/k son separables.

Recíprocamente, si K/F y F/k son separables, sea E la cerradura separable de k en K. Es claro que $F \subseteq E$; y que K es separable sobre E. Por tanto, $K = E$. Por otra parte, $[K : k]_s = 1 \Leftrightarrow [K : F]_s = [F : k]_s = 1$, por lo que $\mathcal{C}$ se satisface en ambos casos.

Para verificar la condición $\mathcal{T}$, supongamos que F_1/k y F_2/k son extensiones tales que $F_1 \subseteq E$ y $F_2 \subseteq E$ para algún campo E.

Supongamos primero que F_1/k es separable y que K es la cerradura separable de F_2 en F_1F_2. Entonces $F_2 \subseteq K$ y todo elemento de F_1 es separable sobre k y sobre F_2, por lo que $F_1 \subseteq K$. Así, $K = F_1F_2$.

Ahora supongamos que $[F_1 : k]_s = 1$. Si $\varphi : F_1F_2 \to \overline{k} = \overline{F_2}$ es un F_2-morfismo, entonces φ es también un k-morfismo cuya restricción a F_1 es única. Se concluye que φ es el único F_2-morfismo de F_1F_2 a $\overline{F_2}$. Así, $[F_1F_2 : F_2]_s = 1$. □

Corolario 3.29 *Una extensión algebraica generada por elementos separables (inseparables puros) es separable (inseparable pura).*

Proposición 3.30 *Para un campo k de característica p las siguientes condiciones son equivalentes:*

a) $k = k^p$.

b) El morfismo de Frobenius es suprayectivo.

c) Toda extensión algebraica de k es separable.

d) Todo elemento algebraico sobre k es separable.

Demostración: Claramente tenemos que *a)* $\Leftrightarrow$ *b)* y que *c)* $\Leftrightarrow$ *d)*. Veamos que *a)* $\Rightarrow$ *d)*: Sea α algebraico sobre k con polinomio mínimo $f(X)$, que es entonces mónico e irreducible en $k[X]$. Si $f(X)$ no es separable, entonces la Proposición 3.25 garantiza que $f(X) = a_0 + a_1X^p + a_2X^{2p} + \cdots + a_rX^{rp}$. Como $k = k^p$, para cada i existe $b_i \in k$ con $a_i = b_i^p$; y entonces el polinomio $f(X) = (b_0+b_1X+b_2X^2+\cdots+b_rX^r)^p$ no es irreducible. Esta contradicción demuestra que α es separable.

Finalmente, veamos que *d)* $\Rightarrow$ *a)*: Si $a \in k$ es tal que $a \notin k^p$, entonces existe $b \in \overline{k}$ tal que $b \notin k$ y $b^p = a$; pero la Proposición 3.27 dice que entonces b es inseparable puro sobre k. Esta es una contradicción. □

Decimos que un campo k es **perfecto** cuando la característica de k es cero, o bien k satisface las condiciones de la proposición anterior. Dejamos de suponer que la característica de k es p.

Corolario 3.31 *Toda extensión algebraica de un campo perfecto es perfecta.*

Demostración: Si F/k es una extensión algebraica y k es un campo perfecto, entonces tendremos para toda extensión algebraica K/F, que K/k es algebraica y separable. Se concluye que K/F es separable. □

Teorema 3.32 *Sea F/k una extensión algebraica con cerradura separable A y cerradura inseparable pura B. Entonces:*
a) F/A es inseparable pura.
b) F/B es separable si F/k es normal.
c) $F = AB$ si F/k es normal.

Demostración: *a*): Sean $\alpha \in F$ y $f(X) = \text{Polmin}(\alpha, k)$. Gracias a la Proposición 3.25, existe un polinomio separable e irreducible $h(Y) \in k[Y]$ tal que $f(X) = h(X^{p^r})$. Esto implica que $\alpha^{p^r} \in A$, es decir, que α es inseparable puro sobre A. Por tanto, F/A es inseparable pura.

b): Sean $\alpha \in F$, $f(X) = \text{Polmin}(\alpha, k)$ y $C = \{\sigma(\alpha) \mid \sigma \text{ es un } k - \text{automorfismo de } F\}$. Como $C \subseteq \{\text{raíces de } f(X)\}$, vemos que C es finito. Además, $\alpha \in C$.

Si $\varphi : F \to \overline{k}$ es un k-morfismo, entonces φ se restringe a un automorfismo de F porque F/k es normal. Además, $\varphi(C) \subseteq C$. Como $\varphi|_C$ es inyectivo, tenemos que $\varphi|_C$ es una biyección.

Sea $g(X) = \prod_{\beta \in C}(X - \beta)$. Entonces $g(\alpha) = 0$ y $\psi(g) = g$ para todo k-morfismo $\psi : F \to \overline{k}$. Esto significa que los coeficientes de g quedan fijos ante todo ψ, por lo que $g(X) \in B[X]$. Esto demuestra que α es separable sobre B. Por tanto, F/B es separable.

c): Con F/k es normal y $\gamma \in F$, tenemos que γ es separable e inseparable puro sobre AB. Así, $\gamma \in AB$. □

Ejercicios

1. Sean k un campo de característica $p > 0$ y F/k una extensión finita. Definimos el **grado de inseparabilidad** de F/k como $[F : k]_i = [F : k]/[F : k]_s$. Demuestre que:

   a) $k \subseteq F \subseteq K \Rightarrow [K : k]_i = [K : F]_i[F : k]_i$;

   b) $[F : k]_i$ es una potencia de p.

   c) Si α es algebraico sobre k, entonces $[k(\alpha) : k]_i$ es la multiplicidad de α en su polinomio mínimo.

2. Sean F/k una extensión finita y F un campo perfecto. Demuestre que k es perfecto.

3. Sea α algebraico sobre un campo k de característica $p > 0$. Demuestre que α es separable sobre k si y sólo si $k(\alpha) = k(\alpha^p)$.

4. Dé un ejemplo de una extensión finita de campos que no sea separable ni inseparable pura.

3.5 Teoría de Galois

Se dice que una extensión de campos F/k es **de Galois** cuando es normal y separable. Como F/k es normal, esto implica que F/k es algebraica.

Dada una extensión arbitraria de campos F/k, definimos su **grupo de Galois**, escrito $\mathrm{Gal}(F/k)$, como el grupo de los k-automorfismos de F. Si suponemos que F/k es de Galois y que $F \subseteq \overline{k}$, entonces podremos identificar $\mathrm{Gal}(F/k)$ con $\{\varphi : F \to \overline{k} \mid \varphi \text{ es un } k-\text{morfismo}\}$.

Iniciamos con un lema importante:

Lema 3.33 (Artin) *Sea $\mathcal{G}$ un grupo finito de automorfismos de un campo F y sea $k = F^{\mathcal{G}} = \{\alpha \in F \mid \sigma(\alpha) = \alpha \text{ para todo } \sigma \in \mathcal{G}\}$. Entonces $[F : k] \leq \circ(\mathcal{G})$.*

Demostración: Digamos que $\mathcal{G} = \{\sigma_1 = 1, \sigma_2, ..., \sigma_n\}$ y supongamos que $A = \{\beta_1, ..., \beta_{n+1}\} \subseteq F$. Veamos que A es linealmente dependiente sobre k.

Consideremos el sistema de n ecuaciones lineales en $n+1$ incógnitas $x_1, ..., x_{n+1}$:

$$\begin{aligned} \sigma_1(\beta_1)x_1 + \cdots + \sigma_1(\beta_{n+1})x_{n+1} &= 0 \\ \sigma_2(\beta_1)x_1 + \cdots + \sigma_2(\beta_{n+1})x_{n+1} &= 0 \\ &\cdots \\ \sigma_n(\beta_1)x_1 + \cdots + \sigma_n(\beta_{n+1})x_{n+1} &= 0 \end{aligned} \tag{3.2}$$

Este sistema admite una solución $(0, ..., 0) \neq (a_1, ..., a_{n+1}) \in F^{n+1}$. Después de reordenar los índices podemos suponer que

$$a_1 = 1, a_2 \neq 0, ..., a_r \neq 0, a_{r+1} = 0, ..., a_{n+1} = 0;$$

y que r es mínimo con esta propiedad.

La primera ecuación puede escribirse como $a_1\beta_1 + \cdots + a_r\beta_r = 0$.

Afirmamos que $a_i \in k$ para todo i, lo que demostrará que A es linealmente dependiente sobre k. Supongamos que esto es falso; y que por ejemplo $a_2 \notin k$. Entonces existe $\tau \in \mathcal{G}$ tal que $\tau(a_2) \neq a_2$. Aplicamos τ al sistema (3.2) para obtener

$$\sum_{j=1}^{n+1} \tau\sigma_i(\beta_j)\tau(a_j) = 0 \quad \text{para} \quad 1 \leq i \leq n. \tag{3.3}$$

Esto dice que $1 = a_1 = \tau(a_1), ..., \tau(a_r), 0, ..., 0$ es otra solución de (3.2), pues $\{\tau\sigma_i\} = \{\sigma_i\}$. De ambas soluciones obtenemos por diferencia la solución $0, a_2 - \tau(a_2) \neq 0, ..., a_r - \tau(a_r), 0, ..., 0$ con menos términos distintos de cero. Esta contradicción demuestra que $[F : k] \leq \circ(\mathcal{G})$. □

Decimos que E es un **campo intermedio** de la extensión F/k cuando $k \subseteq E \subseteq F$.

Teorema 3.34 (Fundamental de la Teoría de Galois) *Sea F/k una extensión finita de Galois con $G = \mathrm{Gal}(F/k)$. Entonces:*

a) La función que a un campo intermedio E le asocia el grupo H de los E-automorfismos de F, es una biyección del conjunto de los campos intermedios al conjunto de los subgrupos de G, cuyo inverso envía cada subgrupo de G al conjunto de sus puntos fijos.

b) La extensión F/E es de Galois para todo campo intermedio E.

c) La extensión E/k es de Galois si y sólo si el subgrupo asociado a E es normal en G. Cuando esto sucede, $\varphi : G \to \mathrm{Gal}(E/k)$ dado por $\varphi(\sigma) = \sigma|_E$ es un morfismo suprayectivo de grupos con núcleo $\mathrm{Gal}(F/E)$, de manera que $\mathrm{Gal}(E/k) \cong G/\mathrm{Gal}(F/E)$.

d) Si el campo intermedio E está asociado al subgrupo H de G y $\sigma \in G$, entonces σE es un campo intermedio asociado al subgrupo $\sigma H \sigma^{-1}$ de G.

e) $[F : k] = \circ(G)$.

Demostración: *e)* Como F/k es separable, $[F : k] = [F : k]_s$. Este último número es el orden de $\{k - \text{morfismos } \psi : F \to \overline{k}\}$, que a su vez coincide con $\circ(G)$ porque F/k es normal y todo k-morfismo ψ se restringe a un k-automorfismo de F.

b) F/E es normal por la Proposición 3.20; y es separable por el Teorema 3.28.

a) Sean E un campo intermedio y $H = \{E - \text{automorfismos de } F\}$. Es claro que $H < G$. Digamos que

$$E' = F^H = \{a \in F \mid \sigma(a) = a, \text{ para todo } \sigma \in H\}.$$

Entonces es inmediato que E' es un campo y que $E \subseteq E' \subseteq F$.

$[E' : E]_s = 1$ porque si $\psi : E' \to \overline{E} = \overline{k}$ es un E-morfismo, entonces ψ se extiende a un E-automorfismo de F, que entonces fija a los elementos de E'; y es por tanto único. Como E'/E es separable, se tiene que $E' = E$.

Hemos demostrado que la función de campos intermedios a subgrupos de G es inyectiva. Veamos ahora que es suprayectiva.

Si $\mathcal{G}$ es un grupo finito de automorfismos de F, entonces $F^{\mathcal{G}}$ es un campo. Si además $\mathcal{G} \subseteq G$, entonces $k \subseteq F^{\mathcal{G}}$. Ahora es suficiente ver que $\mathcal{G} = \mathrm{Gal}(F/F^{\mathcal{G}})$.

Ahora bien, $\mathcal{G} \subseteq \mathrm{Gal}(F/F^{\mathcal{G}})$ es una tautología, de manera que se tiene $\circ(\mathcal{G}) \leq \circ(\mathrm{Gal}\, F/F^{\mathcal{G}}) = [F : F^{\mathcal{G}}]$, esto último en vista de *e)*. La desigualdad $[F : F^{\mathcal{G}}] \leq \circ(\mathcal{G})$ del Lema 3.33 completa la demostración de *a)*.

d) Si E es un campo intermedio y $\sigma \in G$, entonces es claro que σE es otro campo intermedio y que

$$(H = \text{ estabilizador de } E) \Rightarrow (\sigma H \sigma^{-1} \subseteq \text{ estabilizador de } \sigma E).$$

La igualdad se obtiene por simetría.

c) En vista de *d)*, tenemos que $H \lhd G \Leftrightarrow \sigma E = E$ para todo $\sigma \in G$. Pero esto último es equivalente a decir que la extensión E/k es normal, que a su vez es equivalente con E/k de Galois, dado que E/k es separable.

En estas condiciones, todo φ se puede restringir a E, de manera que $\varphi \mapsto \varphi|_E$ es un morfismo suprayectivo de grupos $G \to \mathrm{Gal}(E/k)$, cuyo núcleo es claramente $\mathrm{Gal}(F/E)$. □

Observaciones. Las siguientes afirmaciones son más o menos inmediatas, en ellas usamos la notación del teorema.

1. La correspondencia del teorema voltea inclusiones. Al campo F le corresponde el subgrupo $\{1\}$; mientras que al campo k le corresponde el subgrupo G.

2. Si en la correspondencia del teorema, $E_1 \longleftrightarrow H_1$ y $E_2 \longleftrightarrow H_2$, entonces $E_1E_2 \longleftrightarrow H_1 \cap H_2$ y $E_1 \cap E_2 \longleftrightarrow \langle H_1, H_2\rangle$. Esto se puede verificar directamente o deducir de la observación anterior.

3. Una extensión finita de Galois admite un número finito de campos intermedios. Esto es porque un grupo finito tiene un número finito de subgrupos.

4. Una extensión separable finita E/k admite un número finito de campos intermedios. Esto es porque al tomar una cerradura normal F, obtenemos la extensión finita de Galois F/k, que admite un número finito de campos intermedios.

Teorema 3.35 (Steinitz) *Una extensión finita F/k es simple si y sólo si el número de campos intermedios es finito.*

Demostración: Supongamos que $F = k(\alpha)$ es simple y que $f(X) = \mathrm{Polmin}(\alpha, k)$. Si E es un campo intermedio, entonces consideramos a $g(X) = \mathrm{Polmin}(\alpha, E)$ y llamamos K al campo generado por los coeficientes de $g(X)$ sobre k.

$$\begin{array}{c} F \\ | \\ E \\ \| \\ K \\ | \\ k \end{array}$$

Tenemos que $F = K(\alpha)$ y que $[F : K] \leq \mathrm{gr}\, g(X) = [F : E]$. De aquí se obtiene que $E = K$ está determinado por los coeficientes de $g(X)$. El número de posibilidades para campos intermedios es finito, debido a que $g(X) \mid f(X)$.

Recíprocamente, supongamos que el número de campos intermedios es finito y consideremos dos casos:

Cuando k es finito, F también lo es; y entonces $F^\star$ es cíclico. Todo generador de $F^\star$ genera a F sobre k .

Cuando k es infinito, consideramos solamente el caso $F = k(u, v)$, al cual llegamos por inducción, al ser F finitamente generado sobre k. Como k es infinito mientras que la colección de campos intermedios es finita, se obtienen elementos distintos $b, c \in k$ tales que $k(u + bv) = k(u + cv)$; y entonces

$$v = \frac{(u + bv) - (u + cv)}{b - c} \in k(u + bv) \text{ y también } u \in k(u + bv).$$

Esto implica que $k(u, v) = k(u + bv)$. □

Teorema 3.36 (del Elemento Primitivo) *Toda extensión separable y finita es simple.*

Demostración: Se sigue del Teorema 3.35 y de la Observación 4. □

Teorema 3.37 (Artin) *Sea $\mathcal{G}$ un grupo finito de automorfismos de un campo F y sea $k = F^{\mathcal{G}}$. Entonces F/k es una extensión finita de Galois con* $\mathrm{Gal}(F/k) = \mathcal{G}$.

Demostración: Dado $\alpha \in F$, existe un subconjunto máximo $\{\sigma_1, \ldots, \sigma_n\}$ de $\mathcal{G}$ tal que $\sigma_i(\alpha) \neq \sigma_j(\alpha)$ si $\neq j$. Escribimos $\sigma_i(\alpha) = \alpha_i$, con $\alpha = \alpha_1$.

El polinomio $f(X) = (X - \alpha_1) \cdots (X - \alpha_n)$ es separable, se anula en α, se factoriza totalmente en $F[X]$ y es de grado $n \leq \circ(\mathcal{G})$.

Debido a la maximalidad de $\{\sigma_1, \ldots, \sigma_n\}$, el polinomio $f(X)$ queda fijo ante la acción de $\mathcal{G}$: Para $\gamma \in \mathcal{G}$, el conjunto $\{\gamma\sigma_i(\alpha)\}$ no puede contener nada nuevo. Así, $f(X) \in k[X]$ y F/k es una extensión algebraica separable. Vemos que F es el campo de descomposición de los polinomios mínimos sobre k de todos sus elementos. Así, F/k es normal y de Galois.

Por el Lema 3.33, tenemos que $[F : k] \leq \circ(\mathcal{G})$; pero $[F : k] = \circ(\mathrm{Gal}(F/k))$ y $\mathcal{G} \subseteq \mathrm{Gal}(F/k)$ implican que $\mathcal{G} = \mathrm{Gal}(F/k)$ y que F/k es una extensión finita. □

Teorema 3.38 *Sean E y F extensiones de k contenidas en un campo con E/k finita de Galois. Entonces EF/F y $E/(E \cap F)$ también son finitas de Galois, con grupos de Galois isomorfos.*

Demostración: Las extensiones EF/F y $E/(E \cap F)$ son finitas y de Galois gracias a las Proposiciones 3.8 y 3.20; y al Teorema 3.28. Dado $\sigma \in \mathrm{Gal}(EF/F)$, tenemos que $\sigma|_k = 1$ y que E/k normal $\Rightarrow \sigma(E) = E$. Esto nos permite definir una función $f : \mathrm{Gal}(EF/F) \to \mathrm{Gal}(E/E \cap F)$ así: $f(\sigma) = \sigma|_E$.

Es claro que f es un morfismo de grupos. Supongamos que $\sigma \in \ker f$. Entonces $\sigma|_E = 1$. Como $\sigma|_F = 1$, tenemos que $\sigma = 1$ en EF, de manera que f es inyectivo.

Sea $\mathcal{G} = \mathrm{Im}\, f$, entonces $a \in E^{\mathcal{G}} \Leftrightarrow a \in E \cap F$. Por tanto, el campo intermedio de $E^{\mathcal{G}}$ es $E \cap F$. Del Teorema Fundamental se obtiene que $\mathcal{G} = \mathrm{Gal}(E/E \cap F)$, por lo que f es suprayectivo. □

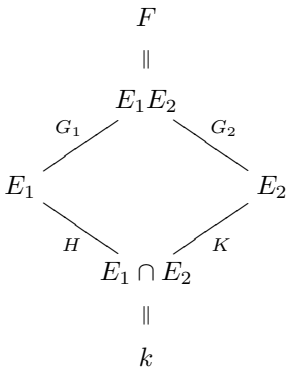
Teorema 3.39 *a) Si F es una extensión finita de Galois de k con grupo de Galois $G \cong G_1 \times G_2$, entonces los campos $E_1 = F^{G_1}$ y $E_2 = F^{G_2}$ satisfacen $F = E_1E_2$ y $k = E_1 \cap E_2$. Además, las extensiones E_1/k y E_2/k son finitas de Galois.*

b) Recíprocamente, si E_1/k y E_2/k son extensiones finitas de Galois con E_1 y E_2 contenidos en un campo, $k = E_1 \cap E_2$, $H = \mathrm{Gal}(E_1/k)$ y $K = \mathrm{Gal}(E_2/k)$, entonces la extensión E_1E_2/k es finita de Galois con grupo $G \cong H \times K$.

Demostración: *a)* Como $G_1, G_2 \lhd G$, tenemos que las extensiones E_1/k y E_2/k son normales; y por tanto, finitas y de Galois.

El subgrupo de G asociado a E_1E_2 es $G_1 \cap G_2 = \{1\}$, así $E_1E_2 = F$.

Similarmente, $E_1 \cap E_2$ corresponde a $G_1G_2 = G$; y por tanto, $E_1 \cap E_2 = k$.



b) El Teorema 3.38 implica que $\mathrm{Gal}(E_1E_2/E_2) \cong \mathrm{Gal}(E_1/E_1 \cap E_2) = H$ y que $\mathrm{Gal}(E_1E_2/E_1) \cong \mathrm{Gal}(E_2/E_1 \cap E_2) = K$. Además, las extensiones E_1 y E_2 de k son normales. Por tanto, E_1E_2/k es una extensión finita de Galois; y también $H, K \lhd G = \mathrm{Gal}(E_1E_2/k)$.

Ahora bien, HK corresponde a $E_1 \cap E_2 = k$ y por eso, $HK = G$. Finalmente, $H \cap K$ está asociado con E_1E_2, por lo que $H \cap K = \{1\}$, que termina la demostración. □

Sean k un campo y $f(X) \in k[X]$ un polinomio separable de grado positivo. Existe un campo de descomposición F de $f(X)$ sobre k, de manera que la extensión F/k es finita de Galois.

Definimos el **grupo de Galois de** $f(X)$ **sobre** k, escrito $\mathrm{Gal}(f/k)$ como $\mathrm{Gal}(F/k)$. Observamos que el isomorfismo de campos de descomposición de un polinomio dado, implica que la definición es correcta, al no depender del campo F.

Cada elemento de $\mathrm{Gal}(f/k)$ queda determinado por su acción en el conjunto de las raíces de $f(X)$. Si $\mathrm{gr}\, f(X) = n$, entonces obtenemos un morfismo inyectivo $\psi : \mathrm{Gal}(f/k) \hookrightarrow S_n$.

Ejemplos.

1. Sean k un campo con caract $k \neq 2$ y $f(X) = aX^2 + bX + c \in k[X]$. Aquí, $F = k(\sqrt{b^2 - 4ac})$ es un campo de descomposición para $f(X)$ sobre k. Tenemos que $f(X)$ es irreducible en $k[X] \Leftrightarrow F \neq k \Leftrightarrow \sqrt{b^2 - 4ac} \notin k \Leftrightarrow [F : k] = 2 \Leftrightarrow \mathrm{Gal}(F/k) \cong S_2 \cong Z_2$.

2. Sean k un campo con caract $k \neq 2$ y $f(X) \in k[X]$ un polinomio separable de grado n con raíces $r_1, ..., r_n \in \overline{k}$, $F = k(r_1, ..., r_n)$ y $G = \mathrm{Gal}(F/k)$, de manera que $G \hookrightarrow S_n$.

   Definimos $\Delta = \prod_{i<j}(r_i - r_j)$, para tener

   $$G \cap A_n = \{\sigma \in G \mid \sigma(\Delta) = \Delta\}, \tag{3.4}$$

   donde A_n es el grupo alternante.

3. Sean k un campo con caract $k \neq 2, 3$ y $f(X) = X^3 + pX + q \in k[X]$ un polinomio separable e irreducible en $k[X]$ con raíces r_1, r_2, r_3; de manera que $F = k(r_1, r_2, r_3)$ es un campo de descomposición de $f(X)$ sobre k.

   Entonces $G = \mathrm{Gal}(f/k) = \mathrm{Gal}(F/k) \hookrightarrow S_3$ actúa transitivamente en $\{r_1, r_2, r_3\}$, por lo que $3 \mid \circ(G)$; y una de dos : $G \cong A_3$ ó bien $G \cong S_3$. Es fácil decidir cual de las dos alternativas ocurre usando el ejemplo anterior y escribiendo $D = \Delta^2$: En el Ejercicio 2.8.1 se pidió demostrar que $D = -4p^3 - 27q^2$, por lo que $G \cong A_3 \Leftrightarrow G \cap A_3 = G \Leftrightarrow \sigma(\Delta) = \Delta$ para todo $\sigma \in G \Leftrightarrow \Delta \in k \Leftrightarrow D \in k^2$.

   El Ejercicio 2.8.5 permite resolver el mismo problema para el caso en que $f(X) = X^3 - a_1X^2 + a_2X - a_3$.

4. Sean k un campo, $R = k[T_1, ..., T_n]$ el anillo de polinomios en n variables sobre k y $F = k(T_1, ..., T_n)$ el campo de fracciones de R. Escribimos los polinomios simétricos elementales así:

   $$s_1 = T_1 + \cdots + T_n,\ s_2 = T_1T_2 + \cdots, ...,\ s_n = T_1 \cdots T_n.$$

   Sea $K = k(s_1, ..., s_n)$. Aquí, F es un campo de descomposición sobre K del polinomio separable

   $$f(X) = \prod_{i=1}^{n}(X - T_i) = X^n - s_1X^{n-1} + \cdots + (-1)^n s_n.$$

   La extensión finita de Galois F/K tiene un grupo de Galois G que deseamos calcular. Sabemos que $G \hookrightarrow S_n$.

   Dada una permutación $\sigma \in S_n$, tenemos una acción natural de σ como k-automorfismo de F así:

   $$\sigma\frac{g(T_1, ..., T_n)}{h(T_1, ..., T_n)} = \frac{g(T_{\sigma(1)}, ..., T_{\sigma(n)})}{h(T_{\sigma(1)}, ..., T_{\sigma(n)})},\ \text{donde } g(T), h(T) \in R.$$

Como $\sigma|_K$ es la identidad, obtenemos un morfismo $S_n \hookrightarrow G$ que nos da $G = S_n$.

Consideremos la cadena $K = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{n-1} = F$, donde $K_i = K_{i-1}(T_i)$, de manera que K_i/K_{i-1} es simple con elemento primitivo T_i, que es raíz del polinomio

$$f_i(X) = \frac{f(X)}{(X - T_1)\cdots(X - T_{i-1})} \in K_{i-1}[X],$$

que es de grado $n - i + 1$. Esto implica que $[K_i : K_{i-1}] \leq n - i + 1$, para $1 \leq i \leq n - 1$. Observe que $f_i(X) \in k[s_1, ..., s_n, T_1, ..., T_{i-1}][X]$. Como

$$\prod_{i=1}^{n-1}(n - i + 1) = n(n-1)\cdots 2 = n! = \circ(S_n) = [F : K],$$

se ve que $[K_i : K_{i-1}] = n - i + 1$, para $1 \leq i \leq n - 1$; y que $f_i(X) = \text{Polmin}(T_i, K_{i-1})$.

Dado un polinomio arbitrario $p(T_1, ..., T_n) \in R$, podemos usar la relación $T_n = s_1 - T_1 - \cdots - T_{n-1}$, para eliminar T_n en favor de las otras T_i y de s_1. Después usamos $f_{n-1}(X)$, que es de grado dos, para eliminar a T_{n-1}^2, expresándolo como polinomio en T_{n-1} de grado ≤ 1 con coeficientes en $k[s_1, ..., s_n, T_1, ..., T_{n-2}]$. Continuamos este proceso hasta expresar a $p(T_1, ..., T_n)$ como combinación lineal de los $n!$ monomios $T_1^{r_1}T_2^{r_2}\cdots T_n^{r_n}$, con $r_i \leq n - i$, para $1 \leq i \leq n$, con coeficientes en $k[s_1, ..., s_n]$. Además, para cada i, el conjunto $\{1, T_i, ..., T_i^{n-i}\}$ es una base de K_i sobre K_{i-1}, por lo que los $n!$ monomios mencionados forman una base de F sobre K. Esto implica que la expresión de $p(T_1, ..., T_n)$ indicada, es única. Observe que en ninguno de estos monomios aparece T_n, por lo que si $p(T_1, ..., T_n)$ es simétrico, entonces solamente aparece el término constante.

Como consecuencia: **Todo polinomio simétrico con coeficientes en un campo, es un polinomio único en los polinomios simétricos elementales**. Compare este resultado con el Teorema 2.60.

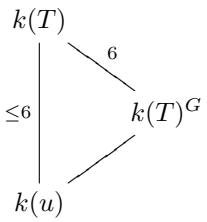
5. Sea G un grupo finito arbitrario. Sabemos que $G \subseteq S_n$ para algún entero n. Usando la notación del ejemplo anterior, sea $E = F^G$. Tenemos la cadena de campos $K \subseteq E \subseteq F$, donde $G = \text{Gal}(F/E)$. Así, vemos que todo grupo finito es el grupo de Galois de una extensión finita de campos. Un caso particular es el del grupo alternante, donde $E = K(\Delta)$, con $\Delta = \prod_{i<j}(T_i - T_j)$, suponiendo que caract $\neq 2$.

6. Sean k un campo con caract $k = 0$ y $k(T)$ el campo de las funciones racionales en una variable sobre k. Supongamos que existe $1 \neq \omega \in k$ tal que $\omega^3 = 1$. Consideremos dos k-automorfismos α y β de $k(T)$ definidos así: $\alpha(T) = \omega T$ y $\beta(T) = T^{-1}$. Sea $G = \langle \alpha, \beta \rangle$. Nos proponemos estudiar G y $k(T)^G$.

   Observamos que α es de orden tres, mientras que β es de orden dos; y que $\beta\alpha\beta^{-1} = \alpha^2$. De manera que G admite la presentación $\langle \alpha, \beta \mid \alpha^3 = \beta^2 = 1,\ \beta\alpha\beta^{-1} = \alpha^2 \rangle$, en la que reconocemos a S_3, o bien descubrimos que existe un isomorfismo de grupos $f : G \to S_3$ tal que $f(\alpha) = (123)$ y $f(\beta) = (12)$.

   Una inspección del escenario nos permite descubrir al elemento $u = T^3 + T^{-3} \in k(T)^G$. La ecuación $T^3 u = T^6 + 1$ implica que la extensión $k(T)/k(u)$ es algebraica con $[k(T) : k(u)] \leq 6$.

   El Teorema de Artin afirma que $k(T)/k(T)^G$ es una extensión de Galois de grado 6 con grupo de Galois G, por lo que tenemos el siguiente diagrama:



   De lo anterior se concluye que $k(T)^G = k(u)$ y que $[k(T) : k(u)] = 6$.

7. Sean p un número primo y $f(X) \in \mathbb{Q}[X]$ un polinomio irreducible con $p-2$ raíces reales y dos raíces complejas conjugadas. Afirmamos que $G = \mathrm{Gal}(f/\mathbb{Q}) \cong S_p$.

   Sabemos que $G \hookrightarrow S_p$; y que G actúa transitivamente en el conjunto de las p raíces de $f(X)$. Esto implica que $p \mid \circ(G)$ y que G contiene un p-ciclo.

   Conjugación compleja estabiliza al conjunto de las raíces de $f(X)$; y por tanto, al campo de descomposición de $f(X)$. Así obtenemos una transposición en G. Como una transposición y un p-ciclo generan a S_p, Ejercicio 1.7.2, se concluye que $G \cong S_p$.

8. Calcularemos $\mathrm{Gal}(f/\mathbb{Q})$ con $f(X) = (X^2-3)(X^2-7)(X^2-17)$. Sean F el campo de descomposición de $f(X)$ sobre $\mathbb{Q}$; y sean E_1, E_2, E_3 los campos de descomposición de X^2-3, X^2-7, X^2-17 respectivamente, también sobre $\mathbb{Q}$.

A partir de la Proposición 3.4, tenemos que $E_1 \cap E_2 = \mathbb{Q}$ y que $E_1E_2 \cap E_3 = \mathbb{Q}$. Como para $1 \leq i \leq 3$, tenemos que $\mathrm{Gal}(E_i/\mathbb{Q}) \cong Z_2$, el Teorema 3.39 nos permite concluir que $\mathrm{Gal}(f/\mathbb{Q}) \cong Z_2 \times Z_2 \times Z_2$.

Ejercicios

1. Describa al grupo $\mathrm{Aut}\,\mathbb{Q}(\sqrt[3]{3})$.
2. Calcule el grupo $\mathrm{Gal}(X^3 - 4X + 2/\mathbb{Q})$.
3. Calcule el grupo $\mathrm{Gal}(X^6 - 2/\mathbb{Q})$.
4. Dada la cadena de campos $k \subset F \subseteq k(X)$, demuestre que la extensión $k(X)/F$ es algebraica.
5. Sean p un número primo y $\alpha, \beta, \gamma \in \mathbb{C}$ tales que
$$\alpha + \beta + \gamma = \alpha\beta + \alpha\gamma + \beta\gamma = \alpha\beta\gamma = p.$$
Demuestre que $\mathbb{Q}(\alpha, \beta, \gamma)$ tiene un automorfismo σ tal que
$$\sigma(\alpha) = \beta,\ \sigma(\beta) = \gamma,\ \sigma(\gamma) = \alpha.$$
6. Sean F/k una extensión finita de Galois con $G = \mathrm{Gal}(F/k)$ y
$$K = \{u \in F \mid \sigma\tau(u) = \tau\sigma(u),\ \forall\ \sigma, \tau \in G\}.$$
Demuestre K es un campo, que la extensión K/k es normal y que $\mathrm{Gal}(K/k)$ es abeliano.

3.6 Campos Reales

Se dice que un campo K es un **campo ordenado** cuando está provisto de un subconjunto K_+ tal que:

1. Dado $a \in K$, se tiene exactamente una de las tres posibilidades: $a \in K_+,\ a = 0,\ -a \in K_+$.
2. $a, b \in K_+ \Rightarrow a + b, ab \in K_+$.

Los elementos de K_+ se llaman **positivos**.
En todo campo ordenado K, se tiene que $1 \in K_+$, pues
$$-1 \in K_+ \Rightarrow (-1)(-1) \in K_+.$$
Además, $1 + \cdots + 1 \in K_+$ para cualquier número de sumandos, por lo que $\mathrm{caract}\, K = 0$.
En un campo ordenado K, se tiene una relación de orden definida así:
$$a < b \quad \Leftrightarrow \quad b - a \in K_+.$$
Se dice que un campo ordenado R es **real cerrado** cuando cumple:

1. $a \in K_+ \Rightarrow$ existe $b \in R$ tal que $b^2 = a$.

2. Todo polinomio en $R[X]$ de grado impar tiene una raíz en R.

Ejemplos. Como prototipos de campos reales cerrados tenemos a $\mathbb{R}$ y a la cerradura algebraica de $\mathbb{Q}$ en $\mathbb{R}$. Un punto de vista a veces conveniente, consiste en pensar que los resultados demostrados en esta sección son para $\mathbb{R}$; y que el enfoque es algebraico, por lo que las propiedades de $\mathbb{R}$ que se usan son las que definen a un campo real cerrado.

Teorema 3.40 *Sean R un campo real cerrado y $C = R(\sqrt{-1})$, entonces C es algebraicamente cerrado.*

Demostración:(Gauss-Artin) Como -1 no es un cuadrado en R, el campo $C = R(\sqrt{-1}) = R[X]/(X^2+1)$ es una extensión de R de grado dos. Escribiendo $i = \sqrt{-1}$, tenemos que $\{1, i\}$ es una base de C como espacio vectorial sobre R.

Afirmamos que todo elemento de C admite una raíz cuadrada en C. Si $a + bi \in C$ con $a, b \in R$, entonces

$$\frac{\sqrt{a^2+b^2}+a}{2}, \frac{\sqrt{a^2+b^2}-a}{2} \geq 0$$

garantizan la existencia de $c, d \in R$ tales que

$$c^2 = \frac{\sqrt{a^2+b^2}+a}{2}, d^2 = \frac{\sqrt{a^2+b^2}-a}{2}.$$

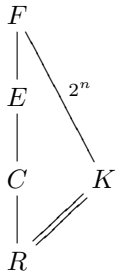
En estas condiciones, $(c+di)^2 = a+bi$, al escoger los signos de c, d adecuadamente.

El Ejemplo 1 de la sección anterior garantiza que C no admite extensiones de grado dos. Veremos ahora que C no admite extensiones finitas propias.

Supongamos que E/C es una extensión finita, entonces también lo son E/R, F/C y F/R, al tomar una cerradura normal F de E/C.

La extensión F/R es entonces finita de Galois. Sean $\mathcal{G} = \text{Gal}(F/R)$, $\mathcal{H}$ un 2-subgrupo de Sylow de $\mathcal{G}$ y $K = F^{\mathcal{H}}$.

Por el Teorema 3.34, tenemos que $[K : R] = \circ(\mathcal{G}) \ / \circ (\mathcal{H})$ es un número impar. Como la extensión K/R es separable; y por tanto, simple, existe $\alpha \in K$, raíz de un polinomio en $R[X]$ irreducible de grado impar, tal que $K = R(\alpha)$. La hipótesis de que R es real cerrado implica que $R = K$.



De lo anterior, concluimos que $\mathcal{G}$ es un 2-grupo; y lo mismo es cierto de su subgrupo $\mathrm{Gal}(F/C)$, que de no ser trivial, admitiría un subgrupo de índice dos, cuyo campo de puntos fijos vendría a ser una extensión de C de grado dos. Así, $F = C$. □

Corolario 3.41 (Teorema Fundamental del Algebra) *El campo* $\mathbb{C}$ *es algebraicamente cerrado.*

Teorema 3.42 (del Valor Intermedio) *Sean* R *un campo real cerrado y* $f(X) \in R[X]$. *Si* $\alpha < \beta \in R$ *son tales que* $f(\alpha)f(\beta) < 0$, *entonces existe* $\gamma \in R$ *con* $\alpha < \gamma < \beta$ *y* $f(\gamma) = 0$.

Demostración: Supongamos que $f(X)$ es mónico. El Teorema 3.40 nos permite saber que la factorización irreducible del polinomio es

$$f(X) = \prod_{i,j}(X - r_i)(X^2 + b_jX + c_j),$$

donde $b_j^2 - 4c_j < 0$ para todo j, por lo que para cualesquiera γ y j se tiene

$$(\gamma^2 + b_j\gamma + c_j) = (\gamma + \frac{b_j}{2})^2 + (c_j - \frac{b_j^2}{4}) > 0.$$

Así, $\beta > \alpha > r_i$ para toda i implica que

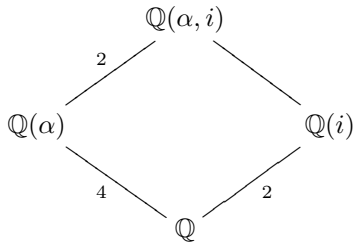
$$f(\alpha)f(\beta) = \prod_{i,j}(\alpha - r_i)(\beta - r_i)(\alpha^2 + b_j\alpha + c_j)(\beta^2 + b_j\beta + c_j) > 0.$$

Similarmente, $\alpha < \beta < r_i$ para toda i, implica que $f(\alpha)f(\beta) > 0$. Se concluye que para algún i, se cumple $\alpha \leq r_i \leq \beta$. □

Ejemplo. Sean p un número primo, $f(X) = X^4 - p \in \mathbb{Q}[X]$ y $G = \mathrm{Gal}(f/\mathbb{Q})$. El polinomio $f(X)$ es irreducible sobre $\mathbb{Q}$ por el Criterio de Eisenstein. Sea α una raíz de $f(X)$, entonces $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 4$.

El conjunto de las raíces de $f(X)$ es $\{\pm\alpha, \pm i\alpha\}$, por ser $\{\pm 1, \pm i\}$ el conjunto de las raíces de $X^4 - 1$, donde $i^2 = -1$. Como $X^2 + 1$ es irreducible, tenemos que $[\mathbb{Q}(i) : \mathbb{Q}] = 2$. La extensión $\mathbb{Q}(i)/\mathbb{Q}$ es de Galois con $\mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \cong Z_2$.

Vemos que $\mathbb{Q}(\alpha, i)$ es un campo de descomposición de $f(X)$. Podemos suponer que α es real, de manera que $i \notin \mathbb{Q}(\alpha)$, por lo que $\mathbb{Q}(\alpha) \cap \mathbb{Q}(i) = \mathbb{Q}$. Así, el Teorema 3.38 nos permite saber que $\mathrm{Gal}(\mathbb{Q}(\alpha, i)/\mathbb{Q}(\alpha)) \cong Z_2$. Además, este grupo está generado por σ, restricción a $\mathbb{Q}(\alpha, i)$ de conjugación compleja.



Tenemos que $\circ(G) = [\mathbb{Q}(\alpha, i) : \mathbb{Q}] = 8$; y que G actúa transitivamente en el conjunto de las raíces de $f(X)$, por lo que existe $\tau \in G$ tal que $\tau(\alpha) = i\alpha$.

- ¿Como actúa τ en i?

- Siendo τ un $\mathbb{Q}$-automorfismo, $\tau(X^2+1) = X^2+1$; y entonces $\tau(i) = \pm i$. En todo caso, τ ó bien $\sigma\tau$ fija al elemento i.

Existe pues $\mu \in \{\tau, \sigma\tau\} \subseteq G$ tal que $\mu(i) = i, \mu(\alpha) = \pm i\alpha$.

El orden de μ es cuatro, $G = \langle \mu, \sigma \rangle$; y G queda descrito por la presentación $\langle \mu, \sigma \mid \mu^4 = \sigma^2 = 1, \sigma\mu\sigma^{-1} = \mu^{-1} \rangle$, por lo que $G \cong D_4$. Ver el Ejercicio 1.12.1.

Si $a_1, a_2, ..., a_n$ es una sucesión de elementos distintos de cero en un campo ordenado R, definimos el **número de cambios de signo de la sucesión** como el número de índices i tales que $a_i a_{i+1} < 0$. El número de cambios de signo de una sucesión arbitraria es el de la subsucesión obtenida al omitir toda ocurrencia de 0.

Existe otra posibilidad para contar el número de cambios de signo de una sucesión $a_1, a_2, ..., a_n$ que contiene ceros, que es diametralmente opuesta a la anterior, le llamaremos el **número aumentado de cambios de signo de la sucesión**, para definirlo supongamos que $a_{k+1} = \cdots = a_{k+r-1} = 0$; pero que $a_k, a_{k+r} \neq 0$. Nuestro valor será el número de cambios de signo de la sucesión obtenida de la sucesion original al reemplazar cada a_{k+j} por $(-1)^j a_k$ para $1 \leq j \leq r-1$. Notemos que este proceso produce el máximo número de cambios de signo posible, para sucesiones obtenidas a partir de la sucesion original, reemplazando las ocurrencias de cero por otros números.

Teorema 3.43 (Budan) *Sean R un campo real cerrado y $f(X) \in R[X]$ con $\mathrm{gr}\, f(X) = n$. Escribimos V_c el número de cambios de signo de la sucesión $f(c), f'(c), ..., f^{(n)}(c)$; y V_c' el número aumentado de cambios de signo de la misma sucesión. Si $a, b \in R$ son tales que $a < b$, $f(a)f(b) \neq 0$ y A es el número de raíces de $f(X)$ en el intervalo abierto (a, b), entonces $V_a - V_b' - A$ es un entero par no negativo.*

Demostración: Por simplicidad, razonaremos como si $R = \mathbb{R}$.

Estudiaremos el comportamiento de V_x para la sucesión

$$f(x), f'(x), ..., f^{(n)}(x) \tag{3.5}$$

a medida que x crece. Este valor permanece constante en cualquier intervalo que no incluya raíces de los polinomios en (3.5).

Estudiemos primero el caso en que α es raíz de $f(X)$ de multiplicidad k, es decir, $f(\alpha) = f'(\alpha) = \cdots = f^{(k-1)}(\alpha) = 0$; pero $f^{(j)}(\alpha) \neq 0$ para $j \geq k$. Existe $\epsilon > 0$ tal que el intervalo $(\alpha - 2\epsilon, \alpha + 2\epsilon)$ no contiene más raíces de polinomios en (3.5) que α.

Afirmamos que en la sucesión $f(\alpha - \epsilon), f'(\alpha - \epsilon), ..., f^{(k)}(\alpha - \epsilon)$, dos números consecutivos siempre tienen signos opuestos. Esto es así porque si alguno de ellos (exceptuando al último) es positivo, el polinomio correspondiente decrece; mientras que si es negativo, el polinomio crece.

Por otra parte, en la sucesión $f(\alpha + \epsilon), f'(\alpha + \epsilon), ..., f^{(k)}(\alpha + \epsilon)$, dos números consecutivos siempre tienen signos iguales, pues si alguno de ellos (exceptuando al último) es positivo, el polinomio correspondiente crece; mientras que si es negativo, el polinomio decrece.

Se concluye que el paso de x por α produce una pérdida de k cambios de signo en (3.5).

Consideremos ahora el caso en que α es raíz de $f^{(r)}(X)$ de multiplicidad k, con $r \geq 1$, es decir, $f^{(r)}(\alpha) = f^{(r+1)}(\alpha) = \cdots = f^{(r+k-1)}(\alpha) = 0$; pero $f^{(r+j)}(\alpha) \neq 0$ para $j \geq k$.

Aquí, el paso de x por α produce una pérdida de k cambios de signo en $f^{(r)}(x), f^{(r+1)}(x), ..., f^{(r+k)}(x)$; pero tal vez se recupere un cambio de signo en $f^{(r-1)}(x), f^{(r)}(x)$. El efecto total es una pérdida de un número par de cambios de signo en (3.5), pues $f^{(r-1)}(x)$ y $f^{(r+k)}(x)$ preservan su signo.

Tenemos demostrado el teorema si a y b no son raíces de $f(X)$, ni de sus derivadas, pues los casos anteriores cubren todas las posibilidades.

Cuando alguno de los números a y b es raíz de alguna derivada de $f(X)$; pero $f(a)f(b) \neq 0$, existe $\epsilon > 0$ tal que el intervalo $(a-2\epsilon, a+2\epsilon)$ no contiene más raíces de $f(X)$ ó sus derivadas que a; y el intervalo $(b - 2\epsilon, b + 2\epsilon)$ tampoco contiene más raíces de $f(X)$ o sus derivadas que b. Aquí, $V_a = V_{a+\epsilon}$ y $V_b' = V_{b-\epsilon}$; mientras que las raíces de $f(X)$ en (a, b) son las mismas que las contenidas en $(a + \epsilon, b - \epsilon)$. □

Corolario 3.44 (Regla de los Signos de Descartes) *Sean R un campo real cerrado y $f(X) = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_{n-s}X^{n-s} \in R[X]$ con $a_n a_{n-s} \neq 0$. Sean A el número de raíces positivas de $f(X)$ y B el número de cambios de signo en $\{a_n, a_{n-1}, ..., a_{n-s}\}$. Entonces $B - A$ es un entero par no negativo.*

Demostración: Dividiendo $f(X)$ entre X^{n-s}, excluimos las raíces iguales a cero, por lo que podemos suponer que $n - s = 0$.

Tomamos $a = 0$ y un número $0 < b \in R$ lo suficientemenete grande, mayor que todas las raíces tanto de $f(X)$ como de sus derivadas, hasta

garantizar que $V_b = V_b' = 0$, donde V_b' es el número aumentado de cambios de signo en $f(b), f'(b), ..., f^{(n)}(b)$. Por otro lado, V_a, que es el número de cambios de signo en $a_0, a_1, ..., a_n$, también lo es para $a_n, a_{n-1}, ..., a_0$. La conclusión es inmediata, usando el Teorema de Budan. □

Estudiaremos a continuación el número de raíces contenidas en un intervalo (a, b) de un campo real cerrado R, para el caso de un polinomio *separable* $f(X) \in R[X]$.

Sea $f(X) \in R[X]$ un polinomio separable. Diremos que una sucesión finita $f(X) = f_0(X), ..., f_s(X)$ de polinomios distintos de cero en $R[X]$ es una **sucesión de Sturm** cuando se cumplan las siguientes condiciones:

1. Dos polinomios consecutivos cualesquiera de la sucesión, no tienen raíces comunes.
2. $f_s(X)$ no tiene raíces en R.
3. Si $f(\alpha) = 0$, entonces el producto $f(x)f_1(x)$ cambia de signo: de negativo a positivo, al paso ascendente de x por α.
4. Si $f_j(\alpha) = 0$, con $1 \leq j < s$, entonces $f_{j-1}(\alpha)f_{j+1}(\alpha) < 0$.

Proposición 3.45 *Dado un polinomio separable $f(X) \in R[X]$, con R real cerrado, la siguiente es una sucesión de Sturm: $f_0(X) = f(X), f_1(X) = f'(X), ..., f_s(X)$, donde $f_{j+1}(X) = -r(X)$ para $j \geq 1$, en caso de que $f_{j-1}(X) = q(X)f_j(X) + r(X)$ exprese el algoritmo euclideano y $r(X) \neq 0$.*

Demostración: De la igualdad de ideales

$$(f_0(X), f_1(X)) = (f_1(X), f_2(X)) = \cdots = (f_s(X)),$$

se obtiene que $f_s(X) = \text{m.c.d.}\{f(X), f'(X)\} \neq 0$ es constante; y que la sucesión cumple las condiciones 1 y 2.

A partir de la definición $f_{j-1}(X) = q(X)f_j(X) - f_{j+1}(X)$, se tiene para cualquier raíz α de $f_j(X)$, que $f_{j-1}(\alpha) = -f_{j+1}(\alpha)$; y la condición 4 se cumple.

Finalmente, si $f(\alpha) = 0$ con $f'(\alpha) > 0$, entonces $f(X) = (X - \alpha)g(X)$, donde $g(\alpha) \neq 0$ y además $f'(X) = (X - \alpha)g'(X) + g(X)$, por lo que $f'(\alpha) = g(\alpha) > 0$. De manera que existe $\epsilon > 0$ tal que $f'(X)g(X) \neq 0$ en $(\alpha - \epsilon, \alpha + \epsilon)$ y entonces $f(X)f'(X) = (X - \alpha)g(X)f'(X)$ cambia de signo: de negativo a positivo, al paso ascendente de x por α. □

La sucesión de la proposición se llama **sucesión de Sturm standard**.

Teorema 3.46 (Sturm) *Sean $f(X) \in R[X]$ separable de grado positivo, R real cerrado y*

$$f(X) = f_0(X), f_1(X), ..., f_s(X) \tag{3.6}$$

una sucesión de Sturm. Para dos elementos $a, b \in R$ tales que $f(a)f(b) \neq 0$, el número de raíces de $f(X)$ en el intervalo (a, b) es $V_a - V_b$, donde V_c es el número de cambios de signo en $f_0(c), f_1(c), ..., f_s(c)$.

Demostración: Estudiemos el comportamiento de V_x a medida que x crece: El valor de V_x es constante en cualquier intervalo que no contenga raíces de polinomios en la sucesión de Sturm.

Supongamos que $f_j(r) = 0$, con $1 \leq j < s$. Entonces $f_{j-1}(r)f_{j+1}(r) < 0$; y además existe $\epsilon > 0$ tal que el intervalo $(r - 2\epsilon, r + 2\epsilon)$ no contiene raíces de $f_{j-1}(X)$ ni de $f_{j+1}(X)$, ni tampoco más raíces de $f_j(X)$ que r. De esta manera, tanto $f_{j-1}(X)$ como $f_{j+1}(X)$ preservan sus signos (opuestos) en $(r - 2\epsilon, r + 2\epsilon)$, por lo que las sucesiones

$$f_{j-1}(r-\epsilon), f_j(r-\epsilon), f_{j+1}(r-\epsilon) \quad \text{y} \quad f_{j-1}(r+\epsilon), f_j(r+\epsilon), f_{j+1}(r+\epsilon)$$

tienen el mismo número de cambios de signo. Así, V_x permanece constante al paso ascendente de x por r.

Supongamos ahora que r es raíz del mismísimo $f(X)$, entonces $f_1(r) \neq 0$; y existe $\epsilon > 0$ tal que el intervalo $(r-2\epsilon, r+2\epsilon)$ no contiene raíces de $f_1(X)$, que por tanto mantiene su signo. Si $f_1(r) > 0$, entonces la condición 3 implica que $f(r-\epsilon) < 0, f(r+\epsilon) > 0$. Así, las sucesiones $f(r-\epsilon), f_1(r-\epsilon)$ y $f(r+\epsilon), f_1(r+\epsilon)$ tienen signos $-+$ y $++$ respectivamente, perdiéndose un cambio de signo al paso ascendente de x por r.

Un estudio similar cuando $f_1(r) < 0$, también conduce a la pérdida de un cambio de signo al paso ascendente de x por r.

La conclusión es que V_x solamenta cambia al paso ascendente de x por cada raíz de $f(X)$, descendiendo cada vez en una unidad. □

Ejemplo. Si definimos los siguientes polinomios:

$$\begin{aligned} f_0(X) &= X^3 + pX + q, \\ f_1(X) &= 3X^2 + p, \\ f_2(X) &= -2pX - 3q, \\ f_3(X) &= -4p^3 - 27q^2, \end{aligned} \tag{3.7}$$

con $p, q \in k$ y $p \neq 0$, entonces tendremos que

$$\begin{aligned} f_0(X) &= \frac{1}{3}Xf_1(X) - \frac{1}{3}f_2(X), \\ f_1(X) &= -(\frac{3}{2p}X + \frac{9q}{4p^2})f_2(X) - \frac{1}{4p^2}f_3(X), \end{aligned} \tag{3.8}$$

por lo que $f_0(X), f_1(X), f_2(X), f_3(X)$ es una sucesión de Sturm.

A partir de un número suficientemenete grande, los signos de la sucesión de Sturm se estabilizan, también son estables estos signos para valores menores a cierto número. Supongamos que $-4p^3-27q^2 > 0$. Entonces $p < 0$ y "$f_0(-\infty), f_1(-\infty), f_2(-\infty), f_3(-\infty)$" tiene signos $-+-+$, mientras que "$f_0(\infty), f_1(\infty), f_2(\infty), f_3(\infty)$" tiene signos $++++$. Así, X^3+pX+q tiene sus tres raíces en k.

Si $-4p^3-27q^2 < 0$, entonces "$f_0(-\infty), f_1(-\infty), f_2(-\infty), f_3(-\infty)$" tiene signos $-+\pm-$, mientras que "$f_0(\infty), f_1(\infty), f_2(\infty), f_3(\infty)$" tiene signos $++\pm-$. Así, $X^3 + pX + q$ tiene exactamente una raíz en k.

Supongamos ahora que $-4p^3 - 27q^2 > 0$ y que $R = \mathbb{R}$. Aquí, $p < 0$ y la substitución $X = \sqrt{-4p/3}\, Y$ produce

$$X^3 + pX + q = \frac{8}{3}(-p)\sqrt{\frac{-p}{3}}\, Y^3 + 2p\sqrt{\frac{-p}{3}}\, Y + q = A(4Y^3 - 3Y - c),$$

$$\text{con} \quad A = -\frac{2}{3}p\sqrt{\frac{-p}{3}} \quad \text{y} \quad c = -\frac{q}{A} = \frac{3q}{2p\sqrt{-p/3}}.$$

Nos concentramos en resolver la ecuación $4Y^3 - 3Y = c$. Como se tiene que

$$c^2 = -\frac{27q^2}{4p^3} < 1 \Leftrightarrow -27q^2 > 4p^3 \Leftrightarrow -4p^3 - 27q^2 > 0,$$

existe un ángulo β tal que $\cos\beta = c$; y la identidad $4cos^3\ \alpha - 3\cos\alpha = \cos 3\alpha$, nos permite resolver trigonométricamente nuestra ecuación cúbica:

$$Y = \cos(\frac{\beta}{3}), \cos(\frac{\beta + 2\pi}{3}), \cos(\frac{\beta + 4\pi}{3}), \text{ para cualquier } \beta \text{ con } \cos\beta = c.$$

Ejercicios

1. Describa al grupo $\operatorname{Aut}\mathbb{R}$.

2. Sea K un campo ordenado. Demuestre que K_+ no está bien ordenado.

3. Sean $f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in K[X]$, con K un campo ordenado; y α una raíz de $f(X)$ en K. Demuestre que α está en el intervalo $(-M, M)$, donde $M = 1 + |a_{n-1}| + \cdots + |a_0|$.

4. Un campo ordenado k es **completo** cuando todo subconjunto de k acotado superiormente admita una mínima cota superior. Demuestre que todo campo ordenado y completo es real cerrado.

5. Dé una demostración directa, por inducción, de la Regla de los Signos de Descartes.

   (Sugerencia: Partiendo de $f(X) \in R[X]$, suponga que $\alpha_1, ..., \alpha_s$ son las raíces positivas de $f(X)$. Entonces $f(X) = (X - \alpha_1)\cdots(X - \alpha_s)g(X)$, donde $g(X) \in R[X]$ no tiene raíces positivas en R. El Teorema del Valor Intermedio y el Ejercicio 3.6.3 garantizan que los coeficientes extremos de $g(X)$ tienen el mismo signo. Esto implica que el número de cambios de signo en la sucesión de coeficientes de $g(X)$ es par. Use este razonamiento para reducir el problema al caso en que $f(X) = (X - \alpha)h(X)$, donde $\alpha \in R_+$ y $h(X) \in R[X]$ satisface la Regla)

3.7 Campos Finitos

Sea K un campo finito. Entonces sabemos que $\operatorname{caract} K = p$ es un número primo; y que el campo primo de K es $\mathbb{Z}/p\mathbb{Z}$.

Proposición 3.47 *El orden de todo campo finito es una potencia de su característica.*

Demostración: Si K es un campo con $\circ(K) = q$ y $\operatorname{caract} K = p$, entonces la dimensión de K como espacio vectorial sobre $\mathbb{Z}/p\mathbb{Z}$ es un entero n. Por tanto, $q = p^n$. □

Si F es un campo con $\operatorname{caract} F = p$, entonces la función $\sigma : F \to F$ dada por $\sigma(x) = x^p$, es un morfismo de anillos, el llamado morfismo de Frobenius. Como $\sigma(1) = 1$ y $\ker \sigma$ es un ideal de F, se tiene que σ es inyectivo; y por tanto, biyectivo en los importantes casos en que F es finito o F es algebraicamente cerrado. En resumen, $\sigma \in \operatorname{Aut} F$ en esos casos.

Observación. Como consecuencia inmediata de que el morfismo de Frobenius es un automorfismo, se tiene que todo campo finito es perfecto.

Teorema 3.48 *Sean p un número primo y n un entero positivo. Escribiendo $q = p^n$, existe un único subcampo K de $\overline{\mathbb{Z}/p\mathbb{Z}}$ de orden q, que es el campo de descomposición de $X^q - X$ sobre $\mathbb{Z}/p\mathbb{Z}$ y que es también el conjunto de las raíces de este polinomio. Todo campo con q elementos es isomorfo a K.*

Demostración: Sea $f(X) = X^q - X$, entonces $f'(X) = qX^{q-1} - 1 = -1$, por lo que $f(X)$ y $f'(X)$ no tienen factores comunes; y $f(X)$ es separable.

El morfismo de Frobenius σ es un automorfismo del campo $\overline{\mathbb{Z}/p\mathbb{Z}}$. Sea K el campo de los puntos fijos de σ^n. Así tenemos que

$$a \in K \Leftrightarrow a^{p^n} = a \Leftrightarrow a \text{ es raíz de } f(X).$$

Si $E \subseteq \overline{\mathbb{Z}/p\mathbb{Z}}$ es otro campo con q elementos, entonces el Teorema de Lagrange aplicado al grupo $E^\star$ implica que para todo $0 \neq b \in E$, se tiene $b^{q-1} = 1$. Como 0 también es raíz de $X^q - X$, vemos que $E = K$.

Finalmente, todo campo F con q elementos, es una extensión finita de $\mathbb{Z}/p\mathbb{Z}$, por lo que existe una copia isomorfa de F dentro de $\overline{\mathbb{Z}/p\mathbb{Z}}$. Esa copia es K. □

En base a este teorema, $\mathbb{F}_q$ representará al campo con q elementos.

Teorema 3.49 *Sea $q = p^n$, con p primo y n un entero positivo. Entonces:*
a) La extensión $\mathbb{F}_q/\mathbb{F}_p$ es finita de Galois.
b) $G = \operatorname{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ es cíclico, generado por el morfismo de Frobenius.
c) $G = \operatorname{Aut}(\mathbb{F}_q)$.

Demostración: *a*) La extensión es separable porque $\mathbb{F}_p$ es perfecto; y es normal porque $\mathbb{F}_q$ es un campo de descomposición.

c) Todo automorfismo de $\mathbb{F}_q$ fija a los elementos del campo primo $\mathbb{F}_p$.

b) $\circ(G) = n$, porque $[\mathbb{F}_q : \mathbb{F}_p] = n$. Claramente, $\sigma \in G$; y $\sigma^r = 1 \Leftrightarrow$ todo elemento de $\mathbb{F}_q$ es raíz de $X^{p^r} - X$. Así, el orden de σ en G es n. □

Se dice que una extensión de Galois es **cíclica**, resp. **Abeliana**, según lo sea su grupo de Galois.

Teorema 3.50 *a) Toda extensión de campos finitos es cíclica de Galois.*

b) Si E/k y F/k son extensiones de campos finitos con $[E : k] = a$ y $[F : k] = b$, entonces $E \subseteq F \Leftrightarrow a \mid b$.

c) Sean k un campo finito y $f(X) \in k[X]$ un polinomio separable con factorización $f(X) = f_1(X) \cdots f_r(X)$, donde cada $f_i(X)$ es irreducible de grado n_i. Entonces $\mathrm{Gal}(f(X)/k)$ *es cíclico de orden* $\mathrm{m.c.m.}\{n_1, ..., n_r\}$ *generado por* $(1, 2, ..., n_1) \cdots (n_1 + \cdots + n_{r-1} + 1, ..., n_1 + \cdots + n_{r-1} + n_r)$*, al ordenar las raíces de $f(X)$ adecuadamente.*

Demostración: *a*) Si $F \supseteq k$ son campos finitos de característica p, entonces $\mathbb{F}_p \subseteq k$, mientras que $F/\mathbb{F}_p$ es una extensión cíclica de Galois, por el Teorema 3.49. Como $\mathrm{Gal}(F/k) < \mathrm{Gal}(F/\mathbb{F}_p)$, tenemos que la extensión F/k es cíclica, además de ser claramente de Galois.

b) Tenemos que $E, F \subseteq \overline{k}$; y que $E \subseteq F \Rightarrow a \mid b$, por el Teorema 3.3. Recíprocamente, si $a \mid b$, entonces escribimos $\circ(k) = q$; y observamos que $(q^a - 1) \mid (q^b - 1)$, por lo que $(X^{q^a-1} - 1) \mid (X^{q^b-1} - 1)$ y entonces $(X^{q^a} - X) \mid (X^{q^b} - X)$. Esto último garantiza que $E \subseteq F$, en vista del Teorema 3.48.

c) Como $\mathrm{Gal}(f(X)/k)$ es cíclico, podemos elegir un generador τ. Las órbitas de τ son los conjuntos de las raíces de los distintos $f_i(X)$, por lo que ordenándolas adecuadamente, podemos representar a τ como la permutación $(1, 2, ..., n_1) \cdots (n_1 + \cdots + n_{r-1} + 1, ..., n_1 + \cdots + n_{r-1} + n_r)$, cuyo orden es claramente $\mathrm{m.c.m.}\{n_1, ..., n_r\}$. □

Teorema 3.51 *Sea $q = p^n$, con p primo y n un entero positivo.*

a) Si $p = 2$, entonces todo elemento de $\mathbb{F}_q$ es un cuadrado.

b) Si $p > 2$, entonces los cuadrados de $\mathbb{F}_q^\star$ forman un subgrupo de índice dos, que es el núcleo del morfismo de grupos $\lambda : \mathbb{F}_q^\star \to \{\pm 1\}$, dado por $\lambda(a) = a^{(q-1)/2}$.

c) Si $p > 2$, entonces $a \in \mathbb{Z}$ es un cuadrado módulo p si y sólo si $a^{(p-1)/2} \equiv 1 \pmod{p}$.

Demostración: *a*) Esto es claro, porque el automorfismo de Frobenius está dado por $\sigma(a) = a^2$.

b) Dado $a \in \mathbb{F}_q^\star$, existe $b \in \overline{\mathbb{F}_q}$ tal que $b^2 = a$. Como $a^{(q-1)/2} = b^{q-1}$ y $(a^{(q-1)/2})^2 = 1$, vemos que $a^{(q-1)/2} = b^{q-1} = \pm 1$. Por otra parte, tenemos que $b \in \mathbb{F}_q \Leftrightarrow b^{q-1} = 1$. La función λ es claramente un morfismo de grupos,

λ es suprayectivo porque $X^{(q-1)/2} - 1$ no tiene $q-1$ raíces; por tanto, el núcleo de λ consiste de los cuadrados de $\mathbb{F}_q^\star$.

c) es consecuencia inmediata de *b*), al tomar $n = 1$. □

Observación. El teorema anterior generaliza y simplifica lo estudiado en la Sección 2.6. En particular, la parte *c*) es el Criterio de Euler (Teorema 2.36).

Teorema 3.52 *Sea $q = p^m$, con p primo y m un entero positivo.*

a) *En $\mathbb{F}_q[X]$ se tiene la factorización*

$$X^{q^n} - X = \prod_{d|n} f_d(X),$$

donde cada $f_d(X)$ es mónico de grado d, irreducible en $\mathbb{F}_q[X]$; y el producto se toma sobre todos esos polinomios.

b) *Sea $u_q(d)$ el número de polinomios mónicos e irreducibles de grado d en $\mathbb{F}_q[X]$. Entonces*

$$q^n = \sum_{d|n} d u_q(d).$$

c) *Si μ es la función de Möbius, se tiene que*

$$n u_q(n) = \sum_{d|n} \mu(d) q^{n/d}.$$

Demostración: *a*) Fijemos un polinomio $f_d(X)$ mónico de grado d e irreducible en $\mathbb{F}_q[X]$. Tenemos que $d \mid n \Leftrightarrow \mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^n} \Leftrightarrow$ toda raíz de $f_d(X)$ pertenece a $\mathbb{F}_{q^n}$, porque toda extensión de campos finitos es normal. Así, $f_d(X) \mid (X^{q^n} - X) \Leftrightarrow d \mid n$. Como $X^{q^n} - X$ es separable, se obtiene la conclusión.

b) Se obtiene comparando los grados de los polinomios en la igualdad *a*).

c) La igualdad *b*) afirma que $\exp_q = (id \times u_q) * 1$, donde $\exp_q$ es la función exponencial con base q, id es la función identidad; y 1 es la función constante con valor 1. Como $1 * \mu = \epsilon$; y ϵ actúa como identidad para el producto convolución, se tiene que $\exp_q * \mu = id \times u_q$, que es nuestra conclusión. □

Ejercicios

1. Sean p un número primo, a y b enteros positivos con $a \mid b$, E y F campos con $\circ(E) = p^a$ y $\circ(F) = p^b$. Exhiba explícitamente un generador de $\mathrm{Gal}(F/E)$.

2. Demuestre que en cualquier campo finito, todo elemento es una suma de dos cuadrados.

3. Sean $\{a_1, ..., a_q\}$ los elementos de $\mathbb{F}_q$, con $q > 3$. Demuestre que

$$\sum_{i<j} a_i a_j = 0; \quad \text{y que} \quad \sum_{i=1}^{q} a_i^2 = 0.$$

4. Sea p un número primo. Demuestre que 3 es un cuadrado en todo campo con p^2 elementos.

5. Demuestre que las raíces de $a_n X^{p^n} + a_{n-1} X^{p^{n-1}} + \cdots + a_0 X \in \mathbb{F}_p[X]$ forman un espacio vectorial sobre $\mathbb{F}_p$, donde p es primo.

6. Sean $K = \mathbb{F}_p(X)$, el campo de funciones racionales sobre $\mathbb{F}_p$ y G el grupo de $\mathbb{F}_p$-automorfismos de K generado por η, donde $\eta(X) = X + 1$. Encuentre K^G y $[K : K^G]$.

3.8 Extensiones Ciclotómicas

En esta sección, estudiaremos al polinomio $f(X) = X^n - 1 \in k[X]$, sobre un campo k. Como $f'(X) = nX^{n-1}$, vemos que $X^n - 1$ es separable en los importantes casos de característica cero o característica p con $p \nmid n$.

Las soluciones de $X^n = 1$ forman un subgrupo multiplicativo finito de $\overline{k}^\star$, que es por tanto, cíclico. Se dice que ζ es una **raíz n-ésima primitiva de la unidad** cuando ζ es un generador de este grupo. Las extensiones de la forma $k(\zeta)/k$ se llaman **ciclotómicas**.

Teorema 3.53 *Sea K un campo de característica cero ó p con $p \nmid n$; y sea ζ una raíz n-ésima primitiva de la unidad. Entonces:*
a) La extensión $K(\zeta)/K$ es finita de Galois.
b) ζ es raíz del polinomio

$$\Phi_n(X) = \prod_{\substack{z^n=1 \\ z \text{ primitiva}}} (X - z).$$

c) $\Phi_n(X) \in k[X]$, donde k es el campo primo; y es mónico.
d) $[K(\zeta) : K] \leq \varphi(n)$, donde φ es la función de Euler.
e) Existe un morfismo inyectivo de grupos $\psi : \mathrm{Gal}(K(\zeta)/K) \to (\mathbb{Z}/n\mathbb{Z})^\star$.
f) $\mathrm{Gal}(K(\zeta)/K)$ es Abeliano.

Demostración: *a*) Claramente esta extensión es finita y separable. Como las raíces de $X^n - 1$ son potencias de ζ, vemos que $K(\zeta)$ es el campo de descomposición de $X^n - 1$, de manera que $K(\zeta)/K$ es normal y de Galois.

b) es claro.

c) Escribiendo $G = \mathrm{Gal}(K(\zeta)/K)$, vemos que z primitiva $\Rightarrow \sigma(z)$ es una raíz primitiva, para todo $\sigma \in G$. Así, es claro que $\sigma\Phi_n(X) = \Phi_n(X)$ para todo $\sigma \in G$; y que $\Phi_n(X) \in K[X]$.

Procediendo inductivamente, a partir de $\Phi_1(X) = X - 1 \in k[X]$ y de

$$\Phi_n(X) = \frac{(X^n - 1)}{\prod_{\substack{d|n \\ d \neq n}} \Phi_d(X)},$$

efectuamos la división. El algoritmo Euclideano produce un cociente y un residuo únicos; el cociente es $\Phi_n(X)$; y el residuo es cero. Como el divisor es mónico y está en $k[X]$, obtenemos nuestra conclusión.

d) Toda raíz primitiva es de la forma ζ^m con $(m, n) = 1$, por lo que $\operatorname{gr} \Phi_n(X) = \varphi(n)$.

e) Definimos ψ como sigue: $\psi(\sigma) = m$, si $\sigma(\zeta) = \zeta^m$. Este es claramente un morfismo inyectivo.

f) es inmediato. □

Teorema 3.54 *Sea $k = \mathbb{Q}$, entonces:*

a) $\Phi_n(X) \in \mathbb{Z}[X]$.

b) *El polinomio $\Phi_n(X)$ es irreducible en $\mathbb{Q}[X]$.*

c) $[\mathbb{Q}(\zeta) : \mathbb{Q}] = \varphi(n)$, *donde φ es la función de Euler.*

d) $\operatorname{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\star$.

Demostración: *a*) El razonamiento para *c*) del teorema anterior también se aplica aquí.

b) Supongamos que $\Phi_n(X) = g(X)h(X)$ en $\mathbb{Z}[X]$; con $g(X)$ mónico, irreducible y de grado positivo.

Para todo primo p tal que $p \nmid n$, se tiene que si z es una raíz n-ésima primitiva de la unidad, entonces z^p también lo es. Toda raíz primitiva se puede obtener a partir de z, después de un número finito de pasos de esta forma. De manera que $g(z) = 0 \Rightarrow g(z^p) = 0$ para todo p con $p \nmid n$ implicaría $\Phi_n(X) = g(X)$.

Supongamos ahora que $g(z^p) \neq 0$ para algún p, entonces $h(z^p) = 0$. Siendo $g(X) = \operatorname{Polmin}(z, \mathbb{Q})$, se infiere que $g(X) \mid h(X^p)$. Reducción $\bmod p$, conduce a $h(X^p) = h(X)^p$; y a $g(X) \mid h(X)^p$, en contradicción a la separabilidad de $\Phi_n(X)$ en $\mathbb{F}_p[X]$.

c) Es consecuencia inmediata de *b*).

d) El morfismo inyectivo ψ de *e*) en el teorema anterior es un isomorfismo, porque los grupos son del mismo orden. □

Teorema 3.55 (Wedderburn) *Todo anillo de división finito es un campo.*

Demostración: Sean F un anillo de división finito y k su centro. Entonces k es un campo finito de orden q; y F es un espacio vectorial sobre k de dimensión n. Tenemos que $\circ(k^\star) = q - 1$; y que $\circ(F^\star) = q^n - 1$.

Nos proponemos escribir la ecuación de clase del grupo $F^\star$, para lo cual observamos que el centralizador $Z(a)$ de un elemento $a \in F^\star$ es un subálgebra de F, que es también un subanillo de división de F. Por tanto, $\circ(Z(a)^\star) = q^d - 1$, donde $d = \dim_k Z(a)$. Además, F es un espacio vectorial sobre $Z(a)$ de dimensión d'; por lo que $n = dd'$.

La ecuación de clase es entonces

$$q^n - 1 = (q-1) + \sum_{\substack{d|n \\ d<n}} \frac{q^n - 1}{q^d - 1}, \tag{3.9}$$

donde la suma se toma sobre las clases de conjugación no triviales de $F^\star$.

Lo que estamos tratando de demostrar es que $n = 1$. Supongamos lo contrario, y obtendremos una contradicción usando (3.9).

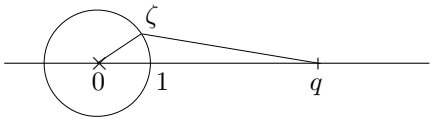
En primer lugar, $[(X^n - 1)/(X^d - 1)] \in \mathbb{Z}[X]$ siempre que $d \mid n$ y que $d < n$. Además, $\Phi_n(X) \mid [(X^n - 1)/(X^d - 1)]$.

Esto implica que $\Phi_n(q) \mid (q^n - 1)/(q^d - 1)$ en $\mathbb{Z}$, para $d \mid n$ con $d < n$, por lo que $\Phi_n(q) \mid (q-1)$.

Esto es absurdo, pues

$$|\Phi_n(q)| = \prod_{\substack{\zeta \text{ primitiva} \\ \zeta^n = 1}} |q - \zeta| > |q - 1|,$$

como se ilustra en la figura. □



$|q - \zeta| > |q - 1|, 1$ para toda ζ

Sea $p > 2$ un número primo. Recordemos que existe un morfismo de grupos $\lambda : \mathbb{F}_p^\star \to \{\pm 1\}$ que define al símbolo de Legendre para todo $a \in \mathbb{F}_p^\star$, o bien para todo entero $a \pmod p$:

$$\lambda(a) = \left(\frac{a}{p}\right) = a^{(p-1)/2}.$$

Definimos la **suma de Gauss** como $S = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right)\zeta^a$, donde ζ es una raíz primitiva p-ésima de la unidad. Esta suma tiene sentido sobre $\mathbb{Q}$, o bien sobre un campo finito $\mathbb{F}_q$ tal que $p \nmid q$.

Teorema 3.56 *La suma de Gauss S definida sobre $\mathbb{Q}$ o sobre un campo finito $\mathbb{F}_q$ con $p \nmid q$ satisface:*

$$S^2 = \left(\frac{-1}{p}\right) p.$$

Demostración: Procedemos a calcular:

$$S^2 = \sum_{a,b=1}^{p-1} \left(\frac{ab}{p}\right)\zeta^{a+b} = \sum_{a,c=1}^{p-1} \left(\frac{a^2c}{p}\right)\zeta^{a(1+c)},$$

al substituir $b = ac$. También tenemos que

$$S^2 = \sum_{c=1}^{p-1} \left(\frac{c}{p}\right) \sum_{a=1}^{p-1} \zeta^{a(1+c)} = \sum_{c=1}^{p-1} \left(\frac{c}{p}\right) \sum_{a=0}^{p-1} \zeta^{a(1+c)},$$

porque $\sum_{c=1}^{p-1} \left(\frac{c}{p}\right) = 0$, ya que exactamente la mitad de los elementos de $\mathbb{F}_p^\star$ son cuadrados. Finalmente, llegamos a

$$S^2 = \left(\frac{-1}{p}\right) \sum_{a=0}^{p-1} 1 + \sum_{c=1}^{p-2} \left(\frac{c}{p}\right) \frac{\zeta^{(1+c)p} - 1}{\zeta^{1+c} - 1} = \left(\frac{-1}{p}\right) p,$$

pues $\zeta^p = 1$. □

Para el caso $p = 2$, tenemos que $(\mathbb{Z}/8\mathbb{Z})^\star = \{\pm 1, \pm 3\}$; y el homomorfismo $\lambda : (\mathbb{Z}/8\mathbb{Z})^\star \to \{\pm 1\}$ dado por $\lambda(\pm 1) = 1$ y $\lambda(\pm 3) = -1$, que nos permite definir la **suma de Gauss modificada** $S_0 = \zeta - \zeta^3 - \zeta^5 + \zeta^7$, donde ζ es una raíz primitiva octava de la unidad. Como

$$\Phi_8(X) = \frac{X^8 - 1}{(X-1)(X+1)(X^2+1)} = \frac{X^8 - 1}{X^4 - 1} = X^4 + 1,$$

tenemos que $1 + \zeta^4 = 0 = \zeta^2 + \zeta^6$; y entonces

$$\zeta^3 + \zeta^5 = \zeta(\zeta^2 + \zeta^4) = \zeta(-\zeta^6 - 1) = -\zeta - \zeta^7.$$

Por tanto, $S_0 = 2(\zeta + \zeta^7)$, que implica $S_0^2 = 4(\zeta^2 + 2 + \zeta^6) = 8$.

Teorema 3.57 *Toda extensión cuadrática de $\mathbb{Q}$ es subciclotómica.*

Demostración: Si $K/\mathbb{Q}$ es una extensión de grado dos, entonces $K = \mathbb{Q}(\sqrt{\epsilon p_1 \cdots p_r})$, donde $\epsilon = \pm 1$ y los p_i son primos distintos.

Aquí, $\zeta_4 = \pm i$, con $i^2 = -1$, mientras que $S_0^2 = 8$, por lo que tenemos $K \subseteq \mathbb{Q}(\zeta_4, \zeta_8, \zeta_{q_1}, ..., \zeta_{q_s})$, donde los q_j son los p_i distintos de dos. □

A continuación, tenemos la **Ley de la Reciprocidad Cuadrática**:

Teorema 3.58 (Gauss) *Si p y q son primos impares distintos, entonces*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Demostración: Sea ζ es una raíz primitiva p-ésima de la unidad sobre $\mathbb{F}_q$. Tenemos que

$$S = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right)\zeta^a \text{ satisface } S^2 = (-1)^{(p-1)/2}p,$$

por lo que calculando en $\mathbb{F}_q$, se tiene que

$$S^{q-1} = [(-1)^{(p-1)/2}p]^{(q-1)/2} = \left(\frac{(-1)^{(p-1)/2}p}{q}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{p}{q}\right),$$

por el Criterio de Euler. Por otro lado,

$$S^q = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right)\zeta^{aq} = \left(\frac{q}{p}\right)S \Rightarrow S^{q-1} = \left(\frac{q}{p}\right).$$

Comparando las dos expresiones para S^{q-1}, obtenemos

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}. \ \square$$

Ejercicios

1. Calcule $\mathrm{Gal}(X^6 + X^3 + 1/\mathbb{Q})$.

2. Sean m y n enteros primos relativos, ζ_m una raíz m-ésima primitiva de la unidad y ζ_n una raíz n-ésima primitiva de la unidad. Demuestre que $(\mathbb{Q}(\zeta_m) \cap \mathbb{Q}(\zeta_n)) = \mathbb{Q}$; y que $\mathbb{Q}(\zeta_m, \zeta_n) = \mathbb{Q}(\zeta_{mn})$.

3. *a*) Sea μ la función de Möbius. Demuestre que

$$\Phi_n(X) = \prod_{d|n}(X^{n/d} - 1)^{\mu(d)}.$$

   b) Encuentre $\Phi_{100}(X),\ \Phi_{360}(X) \in \mathbb{Q}[X]$.

4. Demuestre que para todo primo impar p, se tiene

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

   (Sugerencia: Considere $\beta = \zeta + \zeta^{-1}$, donde $\zeta \in \overline{\mathbb{F}_p}$ es una raíz primitiva octava de la unidad).

3.9 Extensiones Cíclicas

Sean G un grupo y K un campo. Un **carácter** χ de G en K es un morfismo de grupos $\chi : G \to K^\star$. Las funciones de G en K forman un espacio vectorial sobre K que incluye a los caracteres.

Teorema 3.59 (Dedekind) *Toda colección de caracteres distintos de un grupo G en un campo K es linealmente independiente sobre K.*

Demostración: Sea $A = \{\chi_1, ..., \chi_n\}$ una colección de caracteres distintos. Supondremos que A es linealmente dependiente y encontraremos una contradicción. Sea

$$c_1\chi_1 + c_2\chi_2 + \cdots + c_r\chi_r = 0 \tag{3.10}$$

una relación de dependencia con un número mínimo de sumandos, donde los índices se rearreglan, en caso necesario. Aquí, $r \leq n$, todo $c_i \in K^\star$; y claramente $r > 1$.

Como $\chi_1 \neq \chi_2$, existe $h \in G$ tal que $\chi_1(h) \neq \chi_2(h)$. Todo elemento de G se puede escribir como hg con $g \in G$. Por tanto, de (3.10) obtenemos:

$$c_1\chi_1(hg) + c_2\chi_2(hg) + \cdots + c_r\chi_r(hg) = 0;$$

y de ahí

$$c_1\chi_1(h)\chi_1 + c_2\chi_2(h)\chi_2 + \cdots + c_r\chi_r(h)\chi_r = 0. \tag{3.11}$$

Multiplicando (3.10) por $\chi_1(h)$, se tiene

$$c_1\chi_1(h)\chi_1 + c_2\chi_1(h)\chi_2 + \cdots + c_r\chi_1(h)\chi_r = 0. \tag{3.12}$$

Restando (3.12) de (3.11), llegamos a

$$c_2[\chi_2(h) - \chi_1(h)]\chi_2 + \cdots + c_r[\chi_r(h) - \chi_1(h)]\chi_r = 0,$$

que es una relación de dependencia no trivial pues $c_2[\chi_2(h) - \chi_1(h)] \neq 0$; y de longitud menor que la mínima. Esta contradicción nos da el resultado deseado. □

Sea F/k una extensión separable finita, donde $\{\sigma_1, ..., \sigma_n\}$ es el conjunto de k-morfismos de F en $\overline{k}$, de manera que $[F : k] = [F : k]_s = n$. Definimos dos funciones, la **norma** $N_k^F : F \to k$ y la **traza** $Tr_k^F : F \to k$ así:

$$N_k^F(\alpha) = \prod_{i=1}^{n} \sigma_i(\alpha) \quad \text{y} \quad Tr_k^F(\alpha) = \sum_{i=1}^{n} \sigma_i(\alpha).$$

Observaciones. Las siguientes afirmaciones son claras:

1. Como cada σ_i deja fijos a los elementos $N_k^F(\alpha)$, $Tr_k^F(\alpha)$, tenemos que $[N_k^F(\alpha) : k]_s = [Tr_k^F(\alpha) : k]_s = 1 \Rightarrow N_k^F(\alpha),\ Tr_k^F(\alpha) \in k$. De manera que las funciones norma y traza están bien definidas.

2. Si se tiene una cadena de extensiones separables finitas $K \supseteq F \supseteq k$; y $\{\sigma_1, ..., \sigma_n\}$ es el conjunto de k-morfismos de F en $\overline{k}$, los extendemos a morfismos de K en $\overline{k}$ e identificamos $\sigma_1 F \subset \overline{k}$. Sea $\{\tau_1, ..., \tau_m\}$ el conjunto de morfismos de K en $\overline{k}$ que extienden a σ_1. Dado un k-morfismo $\rho : K \to \overline{k}$, existe i tal que $\sigma_i^{-1}\rho|_F = \sigma_1$. Entonces $\sigma_i^{-1}\rho = \tau_j$ para algún j, por lo que $\{\sigma_i\tau_j\}$ es el conjunto de k-morfismos de K en $\overline{k}$. Así, $N_k^F \circ N_F^K = N_k^K$ y también $Tr_k^F \circ Tr_F^K = Tr_k^K$.

3. En el caso de una extensión finita de Galois F/k con $G = \mathrm{Gal}(F/k)$, identificamos a G con el conjunto de k-morfismos de F en $\overline{k}$. Aquí, $N_k^F(\alpha), Tr_k^F(\alpha) \in F^G = k$.

4. $\alpha \in k \Rightarrow N_k^F(\alpha) = \alpha^n, Tr_k^F(\alpha) = n\alpha$.

5. $N(\alpha\beta) = N(\alpha)N(\beta)$, para todos $\alpha, \beta \in F$, por lo que $N_k^F : F^\star \to k^\star$ es un morfismo de grupos.

6. $Tr(\alpha\beta) = Tr(\alpha) + Tr(\beta)$, para todos $\alpha, \beta \in F$; y $Tr_k^F : F_+ \to k_+$ es un morfismo de grupos aditivos.

Sean A un grupo abeliano con operación $+$; y G un grupo multiplicativo que actúa en A. Un **1-cociclo** x es una función $x : G \to A, \sigma \mapsto x_\sigma$ tal que $x_{\sigma\tau} = x_\sigma + \sigma x_\tau$, para todos $\sigma, \tau \in G$. Esto es casi un morfismo de grupos.

Una **1-cofrontera** y es una función $y : G \to A, \sigma \mapsto y_\sigma$ tal que existe $a \in A$ con $y_\sigma = a - \sigma(a)$ para todo $\sigma \in G$.

Proposición 3.60 *a) Los 1-cociclos forman un grupo abeliano $Z^1(G, A)$ ante la suma de funciones.*

b) Toda 1-cofrontera es un 1-cociclo.

c) Las 1-cofronteras son un subgrupo $B^1(G, A)$ de $Z^1(G, A)$.

Demostración: *a)* Si x, y son 1-cociclos, entonces $(x+y)_{\sigma\tau} = x_{\sigma\tau} + y_{\sigma\tau} = x_\sigma + \sigma x_\tau + y_\sigma + \sigma y_\tau = (x+y)_\sigma + \sigma[(x+y)_\tau]$, es decir, $x+y$ es un 1-cociclo.

b) Si x es una 1-cofrontera con $x_\sigma = a - \sigma(a)$ para todo $\sigma \in G$, entonces $x_\sigma + \sigma x_\tau = a - \sigma(a) + \sigma[a - \tau(a)] = a - \sigma\tau(a) = x_{\sigma\tau}$, por lo que x es un 1-cociclo.

c) Si $x_\sigma = a - \sigma(a)$ y $y_\sigma = b - \sigma(b)$ para todo $\sigma \in G$, entonces $(x+y)_\sigma = x_\sigma + y_\sigma = a - \sigma(a) + b - \sigma(b) = (a+b) - \sigma(a+b)$, por lo que $x+y$ es una 1-cofrontera. □

El cociente $H^1(G, A) = Z^1(G, A)/B^1(G, A)$ es el **primer grupo de cohomología** de G en A.

Teorema 3.61 *Sea F/k una extensión finita de Galois con grupo de Galois G. Para la acción de G en $F^\star$ y en F_+, se tiene que*

a) $H^1(G, F^\star) = \{1\}$.

b) $H^1(G, F_+) = (0)$.

Demostración: *a*) Sea x un 1-cociclo (multiplicativo). Tenemos que la función $\sum_{\tau\in G} x_\tau \tau$ no es cero por la independencia lineal de caracteres τ.

Sea $b \in F$ tal que $a = \sum_{\tau\in G} x_\tau \tau(b) \neq 0$. Entonces

$$\sigma(a) = \sum_{\tau\in G} \sigma(x_\tau)\sigma\tau(b) = \sum_{\tau\in G} x_\sigma^{-1} x_{\sigma\tau}\sigma\tau(b) = x_\sigma^{-1} a.$$

Por esto, $x_\sigma = a\sigma(a)^{-1}$, es decir, x es una 1-cofrontera.

b) Debido a la independencia lineal de caracteres, existe $b \in F$ tal que

$$a = \sum_{\sigma\in G} \sigma(b) \neq 0.$$

Reemplazando a b por b/a, obtenemos

$$Tr(b) = \sum_{\sigma\in G} \sigma(b) = 1.$$

Sean x un 1-cociclo (aditivo) y $c = \sum_{\tau\in G} x_\tau \tau(b)$ con b como arriba, entonces para todo $\sigma \in G$ se tiene

$$\sigma(c) = \sum_{\tau\in G} \sigma(x_\tau)\sigma\tau(b) = \sum_{\tau\in G} (-x_\sigma + x_{\sigma\tau})\sigma\tau(b) = -x_\sigma + c.$$

Esto dice que $x_\sigma = c - \sigma(c)$, es decir, que x es una 1-cofrontera. □

Observación. Este es buen momento para notar que si F/k es una extensión finita de Galois, entonces todo elemento de k está en la imagen de Tr: Si $c \in k$ y $\alpha \in F$ es tal que $Tr(\alpha) = 1$, entonces $Tr(c\alpha) = c$.

Teorema 3.62 (Teorema 90 de Hilbert) *Sea F/k una extensión finita de Galois con grupo de Galois $G = \langle\sigma\rangle$ de orden n. Entonces*

a) *$\beta \in F^\star$ tiene norma $1 \Leftrightarrow \beta = \alpha/\sigma(\alpha)$ para algún $\alpha \in F^\star$.*

b) *$\beta \in F$ tiene traza $0 \Leftrightarrow \beta = \alpha - \sigma(\alpha)$ para algún $\alpha \in F$.*

Demostración: *a*) Iniciamos con el cálculo

$$N\left(\frac{\alpha}{\sigma(\alpha)}\right) = \frac{\alpha}{\sigma(\alpha)}\frac{\sigma(\alpha)}{\sigma^2(\alpha)}\cdots\frac{\sigma^{n-1}(\alpha)}{\sigma^n(\alpha)} = 1.$$

Recíprocamente, supongamos que $N(\beta) = 1$, es decir, que

$$\beta\sigma(\beta)\sigma^2(\beta)\cdots\sigma^{n-1}(\beta) = 1.$$

Esto nos permite definir un 1-cociclo así:

$$x_\sigma = \beta, x_{\sigma^2} = \beta\sigma(\beta), ..., x_{\sigma^i} = \beta\sigma(\beta)\cdots\sigma^{i-1}(\beta), ...,$$

$$x_{\sigma^n} = \beta\sigma(\beta)\sigma^2(\beta)\cdots\sigma^{n-1}(\beta) = 1.$$

Para ver que x es realmente un 1-cociclo, supongamos que $\tau = \sigma^i$ y que $\rho = \sigma^j$, entonces

$$x_{\tau\rho} = \beta\sigma(\beta)\cdots\sigma^{i+j-1}(\beta) =$$
$$\beta\sigma(\beta)\cdots\sigma^{i-1}(\beta)\sigma^i[\beta\sigma(\beta)\cdots\sigma^{j-1}(\beta)] = x_\tau\tau(x_\rho).$$

Como $H^1(G, F^\star) = \{1\}$, sabemos que existe $\alpha \in F^\star$ tal que $x_\tau = \alpha\tau(\alpha)^{-1}$ para todo $\tau \in G$. En particular,

$$\beta = x_\sigma = \alpha\sigma(\alpha)^{-1}.$$

b) $Tr[\alpha - \sigma(\alpha)] = [\alpha - \sigma(\alpha)] + [\sigma(\alpha) - \sigma^2(\alpha)] + \cdots + [\sigma^{n-1}(\alpha) - \sigma^n(\alpha)] = 0$. Recíprocamente, si $Tr(\beta) = 0$, definimos un 1-cociclo $y : G \to F_+$ así:

$$y_1 = 0, y_\sigma = \beta, y_{\sigma^2} = \beta + \sigma(\beta), ..., y_{\sigma^{n-1}} = \beta + \sigma(\beta) + \cdots + \sigma^{n-2}(\beta).$$

Ahora bien, $H^1(G, F_+) = (0) \Rightarrow$ existe $\alpha \in F$ con $y_{\sigma^i} = \alpha - \sigma^i(\alpha)$. En particular,

$$\beta = y_\sigma = \alpha - \sigma(\alpha). \square$$

Teorema 3.63 *Sean k un campo y $p \neq \operatorname{caract} k$ un número primo. Supongamos que k contiene una raíz p-ésima primitiva de la unidad ζ.*

a) *Si F/k es una extensión finita y cíclica de Galois de grado p, entonces $F = k(\alpha)$ con α raíz de un polinomio $X^p - a \in k[X]$.*

b) *Si $a \in k$, entonces o bien $X^p - a$ tiene una raíz en k, en cuyo caso se descompone totalmente en $k[X]$, o bien $X^p - a$ es irreducible en $k[X]$; y si α es una raíz, entonces $k(\alpha)$ es un campo de descomposición y $\operatorname{Gal}(k(\alpha)/k)$ es cíclico de orden p.*

Demostración: *a*) Sea $G = \langle\sigma\rangle$. Como $N(\zeta) = 1$, existe $\alpha \in F^\star$ tal que $\zeta = \alpha\sigma(\alpha)^{-1}$; pero entonces $\alpha, \sigma(\alpha) = \zeta^{-1}\alpha, \sigma^2(\alpha) = \zeta^{-2}\alpha, ..., \sigma^{p-1}(\alpha) = \zeta^{-p+1}\alpha$ son p conjugados distintos de α. Por tanto, $[k(\alpha) : k] = p$; y también $F = k(\alpha)$.

Sea $a = \alpha^p$, entonces $\sigma(a) = \sigma(\alpha^p) = \sigma(\alpha)^p = (\zeta^{-1}\alpha)^p = \alpha^p = a$, por lo que $a \in k$.

b) Si $\alpha \in k$ es una raíz de $X^p - a \in k[X]$, entonces el conjunto completo de las raíces es $\{\alpha, \zeta\alpha, ..., \zeta^{p-1}\alpha\} \subseteq k$.

Si $X^p - a$ no tiene raíces en k y α es una raíz en alguna extensión de k, entonces α tiene al menos un conjugado que escribimos $\eta\alpha$ con $\eta \neq 1$. Es claro que $\eta^p = 1$, por lo que η es una raíz primitiva p-ésima de la unidad; y entonces $\eta \in k$. Esto implica que $k(\alpha)$ es un campo de descomposición de $X^p - a$ sobre k.

Sea $G = \operatorname{Gal}(k(\alpha)/k)$. Existe $\sigma \in G$ tal que $\sigma(\alpha) = \eta\alpha$; y es inmediato que σ es de orden p, de manera que $G = \langle\sigma\rangle$ y $X^p - a$ es irreducible. $\square$

Para estudiar el caso en que $p = \operatorname{caract} k$, consideramos al polinomio $X^p - X$. En virtud de la igualdad $(X-1)^p - (X-1) = X^p - X$, se tiene que si α es una raíz, entonces también lo es $\alpha - 1$. Así, el conjunto de raíces de este polinomio es de la forma $\alpha, \alpha+1, \alpha+2, ..., \alpha+(p-1)$.

Teorema 3.64 (Artin-Schreier) *Sean k un campo y $p = \operatorname{caract} k$.*

a) Si F/k es una extensión finita y cíclica de Galois de grado p, entonces $F = k(\alpha)$ con α raíz de un polinomio $X^p - X - a \in k[X]$.

b) Si $a \in k$, entonces o bien $X^p - X - a$ tiene una raíz en k, en cuyo caso se descompone totalmente en $k[X]$, o bien $X^p - X - a$ es irreducible en $k[X]$; y si α es una raíz, entonces $k(\alpha)$ es un campo de descomposición y $\operatorname{Gal}(k(\alpha)/k)$ es cíclico de orden p.

Demostración: *a)* Sea $G = \langle\sigma\rangle$. Como $Tr(1) = 0$, el Teorema 90 de Hilbert produce un elemento $\alpha \in F$ tal que $1 = \alpha - \sigma(\alpha)$; y entonces $\{\sigma(\alpha) = \alpha - 1, \sigma^2(\alpha) = \alpha - 2, ..., \sigma^p(\alpha) = \alpha\}$ tiene p elementos. Por tanto, $[k(\alpha) : k] = p$; y también $F = k(\alpha)$. Además,

$$a = \alpha(\alpha - 1) \cdots [\alpha - (p - 1)] \in F^\sigma = k.$$

Como $\alpha^p - \alpha = a$, vemos que α es raíz de $X^p - X - a \in k[X]$.

b) se puede demostrar como en el teorema anterior. □

Se usa la notación $\wp(\alpha) = \alpha^p - \alpha$; de manera que en el caso del teorema, $\wp^{-1}(a) = \alpha$.

Ejercicios

1. Sea F/k una extensión finita y separable, tal que $F = k(\alpha)$ con $\operatorname{Polmin}(\alpha, k) = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0$. Demuestre que $N_k^F(\alpha) = (-1)^n c_0$ y que $Tr_k^F(\alpha) = -c_{n-1}$.

2. Sea F/k una extensión de campos finitos. Demuestre que la norma $N_k^F : F^\star \to k^\star$ es suprayectiva.

3. Demuestre que $0 \neq a \in \mathbb{Q}$ con $a = b/c$, $(b, c) = 1$ está en la imagen de la norma $N_{\mathbb{Q}}^{\mathbb{Q}(\sqrt{-1})} \Leftrightarrow$ (si la máxima potencia de p que divide ó bien a b ó bien a c es impar, entonces $p \equiv 1 \pmod 4$).

4. Un grupo G es de **exponente** d cuando todo elemento $a \in G$ satisface $a^d = 1$. Sea F el campo de descomposición del conjunto de polinomios $\{X^n - a_1, X^n - a_2, ..., X^n - a_m\}$ sobre k, donde k contiene a las raíces n-ésimas de la unidad y $\operatorname{caract} k \nmid n$. Demuestre que $\operatorname{Gal}(F/k)$ es un grupo abeliano de exponente n.

5. Sea F/k una extensión finita de Galois tal que k contiene a las raíces n-ésimas de la unidad, $\operatorname{caract} k \nmid n$ y $\operatorname{Gal}(F/k)$ es un grupo abeliano de exponente n. Demuestre que F el campo de descomposición de un conjunto de polinomios $\{X^n - a_1, X^n - a_2, ..., X^n - a_m\}$ sobre k.

3.10 Solubilidad con Radicales

Se dice que una extensión separable finita de campos F/k es **soluble con radicales** cuando existe una extensión finita de Galois E/k tal que $F \subseteq E$; y E se obtiene a partir de k por medio de una sucesión finita de extensiones de los siguientes tipos:

1. Adjuntando raíces de ecuaciones de la forma $X^n - 1$ con $p \nmid n$, donde $p = \operatorname{caract} k$.
2. Adjuntando raíces de ecuaciones de la forma $X^p - a$ con p primo, donde $p \neq \operatorname{caract} k$, cuando ya se tienen las raíces p-ésimas de la unidad.
3. Adjuntando raíces de ecuaciones de la forma $X^p - X - a$, donde $p = \operatorname{caract} k$.

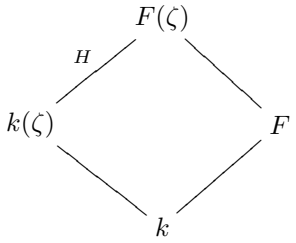
Teorema 3.65 *Sean F/k una extensión finita de Galois y $G = \operatorname{Gal}(F/k)$. Entonces F/k es soluble con radicales si y sólo si G es soluble.*

Demostración: Las extensiones de los tipos 2) ó 3), son cíclicas de orden p por los Teoremas 3.63 y 3.64, en presencia de las raíces de la unidad adecuadas. Toda extensión de tipo 1) tiene grupo de Galois abeliano por el Teorema 3.53 f). Esto implica que si $H = \operatorname{Gal}(E/k)$ con E obtenido a partir de k a través de extensiones de estos tipos, entonces H admite una sucesión de subgrupos subnormales $H = H_0 \supseteq H_1 \supseteq \cdots \supseteq H_r = \{1\}$, tal que todo cociente H_i/H_{i+1} es abeliano. Así, H es soluble. Supongamos que F/k es soluble con radicales, con E y H como en la definición de solubilidad con radicales. Entonces G es soluble al ser imagen homomorfa de H.

Recíprocamente, si G es soluble (y finito), existe una sucesión de subgrupos subnormales $G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_s = \{1\}$, tal que cada cociente G_i/G_{i+1} es cíclico de orden primo p_i.

La cadena de campos correspondiente: $k = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_s = F$, es tal que cada extensión F_{i+1}/F_i es de Galois con grupo cíclico de orden p_i, por lo que F_1/k es de tipo 1), 2) ó 3). El tipo 2) sólo puede ocurrir cuando k contenga a las raíces p_1-ésimas de la unidad; y la demostración concluye por inducción. En caso contrario, adjuntamos una raíz m-ésima primitiva de la unidad ζ al campo k, donde $m = \prod_{p_i \neq \operatorname{caract}} p_i$.

Esta situación está representada en el diagrama



Aquí , $k(\zeta)/k$ es una extensión abeliana por el Teorema 3.53 f), mientras que $H = \mathrm{Gal}(F(\zeta)/k(\zeta)) < G$, gracias al Teorema 3.38 , por lo que H es soluble. En la cadena $k \subseteq k(\zeta) \subseteq F(\zeta)$ se ve que $F(\zeta)$ es soluble con radicales, por lo que F también lo es. □

Se dice que el polinomio separable $f(X)$ con coeficientes en el campo k, es **soluble con radicales** cuando el campo de descomposición de $f(X)$ es una extensión de k soluble con radicales.

Corolario 3.66 (Abel) *La ecuación general de grado n es soluble con radicales si y sólo si $n \leq 4$.*

Demostración: En el Ejemplo 4 de la Sección 3.5, vimos que el grupo de Galois de un polinomio genérico de grado n es S_n, que es soluble exactamenete para $n \leq 4$. □

La ecuación cuadrática general. Si $Y^2 + bY + c \in k[Y]$, $\mathrm{caract}\, k \neq 2$, entonces la substitución

$$Y = X - \frac{b}{2} \text{ produce } X^2 - \left(\frac{b^2}{4} - c\right).$$

Si $\mathrm{caract}\, k = 2$, entonces la substitución $Y = bX$ produce

$$b^2X^2 + b^2X + c = b^2\left(X^2 + X + \frac{c}{b^2}\right) = b^2\left(X^2 - X - \frac{c}{b^2}\right),$$

que se resuelve así: $x_1 = \wp^{-1}(c/b^2),\ x_2 = x_1 + 1$.

La ecuación cúbica general. Según el Ejemplo 3 de la Sección 3.5, el polinomio $f(Y) = Y^3 + aY^2 + bY + c \in k(a,b,c)[Y]$ tiene grupo de Galois S_3, donde $k(a,b,c)$ es el campo de funciones racionales en las variables a, b, c sobre el campo k. Por tanto, la ecuación $f(Y) = 0$ es soluble con radicales. Esto implica que sus raíces pertenecen a un campo que proviene de $k(a,b,c)$ por medio de extensiones de tipos 1), 2) y 3).

Suponemos que $\mathrm{caract}\, k \neq 2, 3$ y adjuntamos a k una raíz cúbica primitiva de la unidad ω para tener que las raíces de $f(Y)$ están en un campo que proviene de $k(a,b,c,\omega)$ por medio de extensiones de tipo 2).

Es conveniente efectuar la substitución $Y = X - \frac{1}{3}a$ para simplificar: $Y^3 + aY^2 + bY + c = X^3 + pX + q$, donde $p = -\frac{1}{3}a^2 + b$ y también $q = \frac{2}{27}a^3 - \frac{1}{3}ab + c$. Escribimos $K = k(p,q,\omega)$.

A la cadena de grupos $S_3 \supset A_3 \supset \{1\}$ le corresponde la cadena de campos $K \subset K(\Delta) \subset K(\Delta,\alpha)$, donde $\Delta = \sqrt{D}$, con $D = -4p^3 - 27q^2$, que es el discriminante de $X^3 + pX + q$; y donde $\alpha^3 \in K(\Delta)$.

Resolver la ecuación $X^3 + pX + q = 0$ significa expresar concretamente sus raíces x_1, x_2, x_3 como elementos de $K(\Delta,\alpha)$, donde α^3 también debe escribirse como elemento de $K(\Delta)$.

Como existe un generador σ de $\mathrm{Gal}(K(\Delta,\alpha)/K(\Delta)) \cong A_3$, que satisface $\sigma(x_1) = x_2, \sigma(x_2) = x_3, \sigma(x_3) = x_1$, proponemos $\alpha = x_1 + \omega x_2 + \omega^2 x_3$, la **resolvente de Lagrange**, basados en que $\sigma(\alpha) = \omega^{-1}\alpha$; y calculamos:

$$\begin{aligned}\alpha^3 = x_1^3 + x_2^3 + x_3^3 + 3\omega(x_1^2x_2 + x_2^2x_3 + x_3^2x_1)\\ +3\omega^2(x_1x_2^2 + x_2x_3^2 + x_3x_1^2) + 6x_1x_2x_3.\end{aligned} \tag{3.13}$$

Por otra parte, $x_1 + x_2 + x_3 = 0$ implica

$$\begin{aligned}0 = x_1^3 + x_2^3 + x_3^3 + 3(x_1^2x_2 + x_2^2x_3 + x_3^2x_1)\\ +3(x_1x_2^2 + x_2x_3^2 + x_3x_1^2) + 6x_1x_2x_3.\end{aligned} \tag{3.14}$$

Restando (3.14) de (3.13), se tiene

$$\alpha^3 = (3\omega - 3)(x_1^2x_2 + x_2^2x_3 + x_3^2x_1) + (3\omega^2 - 3)(x_1x_2^2 + x_2x_3^2 + x_3x_1^2).$$

Para simplificar estas expresiones usamos las igualdades

$$\begin{aligned}(x_1^2x_2 + x_2^2x_3 + x_3^2x_1) - (x_1x_2^2 + x_2x_3^2 + x_3x_1^2) =\\ (x_1 - x_2)(x_2 - x_3)(x_1 - x_3) = \Delta = \sqrt{-4p^3 - 27q^2},\end{aligned}$$

$$\begin{aligned}(x_1^2x_2 + x_2^2x_3 + x_3^2x_1) + (x_1x_2^2 + x_2x_3^2 + x_3x_1^2) =\\ (x_1 + x_2 + x_3)(x_1x_2 + x_2x_3 + x_1x_3) - 3x_1x_2x_3 = 3q.\end{aligned}$$

Así, llegamos a ver que

$$\begin{aligned}\alpha^3 = \frac{3\omega - 3}{2}(3q + \sqrt{D}) + \frac{3\omega^2 - 3}{2}(3q - \sqrt{D}) =\\ -\frac{27}{2}q + \frac{3}{2}\sqrt{-3D},\end{aligned}$$

para lo que usamos las igualdades $\omega^2 + \omega + 1 = 0$ y $\omega - \omega^2 = \sqrt{-3}$. También decidimos desplazar a ω en favor de $\sqrt{-3}$.

Ahora definimos $\beta = x_1 + \omega^2 x_2 + \omega x_3$, que es otra resolvente de Lagrange; y que proviene de α al intercambiar ω con ω^2, por lo que es inmediato que

$$\beta^3 = -\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}.$$

En resumen, tenemos el siguiente sistema de ecuaciones lineales

$$\begin{aligned}x_1 + \ \ x_2 \ + \ \ x_3 &= 0\\ x_1 + \omega x_2 + \omega^2 x_3 &= \sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\sqrt{-3D}}\\ x_1 + \omega^2 x_2 + \omega x_3 &= \sqrt[3]{-\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}}\end{aligned}$$

Este sistema se resuelve, pues el determinante de los coeficientes no es cero, donde la extracción de una raíz cúbica introduce cierta ambigüedad

menor; pero la presencia de dos raíces cúbicas simultáneas presenta una ambigüedad mayor, que para desaparecer requiere de la condición: $\alpha\beta = (x_1+x_2+x_3)^2-3(x_1x_2+x_2x_3+x_1x_3) = -3p$, que también puede escribirse

$$\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3D}}\ \ \sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3D}} = -3p. \tag{3.15}$$

En cuanto a las raíces de $X^3+pX+q=0$, estas vienen dadas por las **fórmulas de Tartaglia-Cardano**:

$$\begin{aligned}
x_1 &= \frac{1}{3}\left[\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3D}}+\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3D}}\right]\\
x_2 &= \frac{1}{3}\left[\omega^2\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3D}}+\omega\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3D}}\right]\\
x_3 &= \frac{1}{3}\left[\omega\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3D}}+\omega^2\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3D}}\right]
\end{aligned} \tag{3.16}$$

sujetas a la condición (3.15).

Proposición 3.67 *Sea $f(X) \in \mathbb{Q}[X]$ un polinomio cúbico e irreducible con tres raíces reales. Entonces la ecuación $f(X)=0$ no se puede resolver por medio de una sucesión de extensiones reales simples.*

Demostración: Supongamos que la afirmación es falsa, que F es el campo de descomposición de $f(X)$, que $\mathbb{Q} = K_0 \subseteq K_1 \subseteq \cdots K_{r-1} \subseteq K_r \subseteq \mathbb{R}$ es una cadena de extensiones reales simples con todo $[K_{i+1} : K_i]$ primo, $F \subseteq K_r$, que $K_1 = K_0(\sqrt{D})$; y que $F \nsubseteq K_{r-1}$.

En estas condiciones, $\mathrm{Gal}(f(X)/K_{r-1}) \cong A_3$, por lo que $f(X)$ es irreducible en K_{r-1}; pero se factoriza totalmente en K_r. Así, $K_r = K_{r-1}(\sqrt[p]{a})$ con p primo; pero $3 \mid p \Rightarrow p = 3$. Además, K_r es un campo de descomposición de $f(X)$ sobre K_{r-1}; pero entonces $\sqrt[3]{a}, \omega\sqrt[3]{a} \in K_r \Rightarrow \omega \in K_r$, en contradicción con $\omega \notin \mathbb{R}$. □

Un ejemplo de Teoría de Grupos. Aquí demostraremos para p primo, que un subgrupo transitivo G de S_p es soluble si y sólo si todo elemento de $G \setminus \{1\}$ tiene cuando más un punto fijo.

Sea G un subgrupo transitivo de S_p, donde p es un número primo.

Paso 1 Todo subgrupo normal N no trivial de G es transitivo.

Demostración: Consideremos la acción de N en $P = \{1, 2, ..., p\}$. La órbita de q_0 tiene $[N : S(q_0)]$ elementos, donde $S(q_0)$ es el estabilizador en N de q_0. Para cualquier otro punto $q_1 \in P$, existe $x \in G$ tal que $xq_0 = q_1$, por lo que $S(q_1) = xS(q_0)x^{-1} \cap N = xS(q_0)x^{-1}$, dada la normalidad de N. Así, todas las órbitas de N tienen el mismo número de elementos, número que divide a p; y por tanto coincide con p, al no ser N trivial. □

Identifiquemos al conjunto P con $\mathbb{F}_p$ y digamos que $\sigma \in S_p$ es **afín** cuando existan $b \in \mathbb{F}_p$ y $a \in \mathbb{F}_p^\star$ tales que $\sigma(x) = ax + b$ para todo $x \in P$. Las **translaciones** son las transformaciones afines con $a = 1$.

Paso 2 Las translaciones $\neq 1$ no tienen puntos fijos. Las transformaciones afines restantes tienen exactamente un punto fijo.

Demostración: La ecuación $ax + b = x$ tiene cuando más la solución $x = -(a-1)^{-1}b$, para lo que se requiere que $a \neq 1$; exceptuando el caso en que $a = 1$ y $b = 0$, que corresponde a la función identidad. □

Paso 3 Si además, G es soluble, existe una sucesión de subgrupos

$$G = G_0 \supset G_1 \supset \cdots \supset G_r \supset G_{r+1} = \{1\}, \tag{3.17}$$

donde cada subrupo es normal en el que le precede, cada cociente es abeliano no trivial y G_r es cíclico de orden p.

Demostración: Tal sucesión subnormal existe por la solubilidad de G, con G_r abeliano; pero se puede conseguir G_r cíclico de orden p, porque todos los subgrupos en ella son transitivos; y por tanto de orden divisible por p, gracias al Paso 1. □

Reordenamos los elementos de P para poder escribir $G_r = \langle\sigma\rangle$, con $\sigma = (12\cdots p)$.

Paso 4 Para G soluble, todo p-ciclo de G está en G_r; y todo elemento de G fuera de G_r tiene exactamente un punto fijo.

Demostración: Veremos por inducción, al ascender en la sucesión 3.17 desde $\{1\}$, que todo elemento de G es afín y que todo p-ciclo de G está en G_r.

Supongamos que ambas afirmaciones son ciertas para G_{r-m}.

Sea $\tau \in G_{r-m-1}, \tau \notin G_r$. Entonces $\tau\sigma\tau^{-1} \in G_{r-m}$ es un p-ciclo, por lo que $\tau\sigma\tau^{-1} \in G_r$ y así $\tau\sigma\tau^{-1} = \sigma^a$ para algún $1 \neq a \in \mathbb{F}_p^\star$, pues cualquier p-ciclo en S_p conmuta solamente con sus potencias.

Supongamos que $\tau(i) = j$; y calculemos:

$$\tau\sigma\tau^{-1}(j) = \sigma^a(j) = j + a \Rightarrow \tau(i+1) = \tau\sigma(i) = \tau(i) + a.$$

Escribiendo $\tau(0) = b$, obtenemos

$$\tau(1) = a + b, \tau(2) = 2a + b, ..., \tau(i) = ia + b, ...$$

por lo que τ es afín. Finalmente, todo p-ciclo de G_{r-m-1}, siendo afín y sin puntos fijos, es una translación, que está en G_r. □

Paso 5 Si G es soluble y transitivo, entonces cada elemento de $G \setminus \{1\}$ tiene cuando más un punto fijo.

Paso 6 Si G es transitivo y todo elemento de $G \setminus \{1\}$ tiene cuando más un punto fijo, entonces G tiene un subgrupo normal de orden p.

Demostración: Claramente, $p \mid \circ(G)$. Sea T un subgrupo de G de orden p. La transitividad también implica que al calcular el número de órbitas de G según Burnside, se tiene

$$\circ(G) = \sum_{g \in G} F_g, \tag{3.18}$$

donde F_g es el número de puntos fijos de g. En la ecuación (3.18), $F_1 = p$, mientras que $F_g = 0$ para todo $1 \neq g \in T$. Como $F_g \leq 1$ para todo $g \in G$, se infiere que $F_g = 1$ para todo $g \notin T$. Así, $g \notin T \Rightarrow g$ no es un p-ciclo; y T contiene a todos los p-ciclos de G. Esto garantiza que T es normal. □

Paso 7 Si G es transitivo y todo elemento de $G \setminus \{1\}$ tiene cuando más un punto fijo, entonces G es soluble.

Demostración: Sea $T = \langle \sigma \rangle \lhd G$ de orden p. La acción obtenida por conjugación de T con los elementos de G da origen a un morfismo $\psi : G \to \mathbb{F}_p^\star$ definido así: Si $x\sigma x^{-1} = \sigma^n$, escribimos $\psi(x) = n$. Aquí, $\ker \psi = T$, pues los p-ciclos conmutan solamente con sus potencias.

El morfismo ψ induce otro morfismo inyectivo $G/T \hookrightarrow \mathbb{F}_p^\star$, que implica la solubilidad de G/T, de manera que G también es soluble. □

Ejercicios

1. Sea k un campo con característica distinta de 2 y de 3. Demuestre que el polinomio $X^3 + pX + q \in k(p,q)[X]$ con raíces x_1, x_2, x_3 admite como campo de descomposición a $k(\sqrt{-4p^3 - 27q^2}, x_1)$.

2. Sea $X^p - a \in \mathbb{Q}[X]$ irreducible con p primo. Demuestre que entonces $\mathrm{Gal}(X^p - a/\mathbb{Q})$ es isomorfo con el grupo $\mathbb{A}_2(\mathbb{F}_p)$ de transformaciones de $\mathbb{F}_p$ de la forma $y \mapsto \alpha y + \beta$, donde $\alpha, \beta \in \mathbb{F}_p$ y $\alpha \neq 0$.

3. **(Galois)** Sean k un campo de característica cero, $f(X) \in k[X]$ irreducible de grado primo p, F un campo de descomposición de $f(X)$ sobre k y $G = \mathrm{Gal}(f(X)/k)$. Demuestre que G es soluble si y sólo si F está generado sobre k por cualquier pareja de raíces de $f(X)$. (Sugerencia: Use el Ejemplo de Grupos).

4. Sean k un subcampo de $\mathbb{R}$, $f(X) \in k[X]$ irreducible y soluble de grado primo. Demuestre que o bien $f(X)$ tiene exactamente una raíz real, o bien todas sus raíces son reales.

3.11 Constructibilidad con Regla y Compás

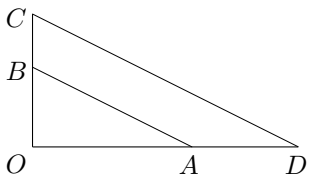
Hay toda una serie de problemas geométricos clásicos que tratan de determinar las construcciones que pueden lograrse con una regla y un compás. Aquí suponemos que una regla no tiene marcadas subdivisiones; y que solamente sirve para trazar la recta que pase por dos puntos distintos dados.

El punto de partida es una longitud identificada con la unidad. Las longitudes que se pueden lograr se llaman **constructibles**, las cuales forman un campo:

Es claro como obtener $\alpha+\beta$ y $\alpha-\beta$ a partir de α y de β. Recordemos como encontrar $\alpha\beta$, y también $1/\beta$ para $\beta\neq 0$.

Dados α y β, construimos un triángulo (rectángulo) $\triangle OAB$, en el que $\overline{OA}=\alpha$ y $\overline{OB}=1$. Prolongamos $\overline{OB}$ hasta $\overline{OC}=\beta$; y trazamos CD paralela a AB. Por la similaridad de los triángulos, se tiene

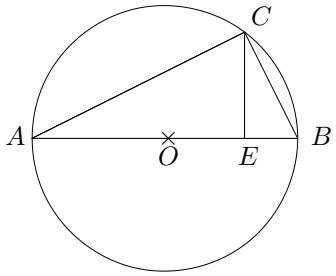
$$\frac{\overline{OA}}{\overline{OB}}=\frac{\overline{OD}}{\overline{OC}},\ \text{es decir}\ \overline{OD}=\frac{\overline{OA}\times\overline{OC}}{\overline{OB}}=\alpha\beta.$$



Repetimos la construcción a partir de $\overline{OA}=1=\overline{OC}$ y de $\overline{OB}=\beta$ para tener $\overline{OD}=1/\beta$.

Es posible construir $\sqrt{d}$ para $d>0$ dado. Esto se hace así:

Trazamos la recta AB de longitud $d+1$ con $\overline{AE}=d$ y $\overline{EB}=1$, después construimos un círculo de diámetro $\overline{AB}$.



Trazamos $CE\perp AB$, entonces $\measuredangle ACB$ es recto, por lo que $\measuredangle EAC=\measuredangle ECB$. De esta manera, $\triangle EAC$ y $\triangle ECB$ son similares; y tenemos que

$$\frac{\overline{AE}}{\overline{EC}}=\frac{\overline{EC}}{\overline{EB}},\ \text{es decir,}\ \overline{EC}^2=\overline{AE}\times\overline{EB}=d,\ \text{ó bien,}\ \overline{EC}=\sqrt{d}.$$

Hemos visto que las longitudes o números reales constructibles forman un subcampo de $\mathbb{R}$ que contiene al campo primo $\mathbb{Q}$ y a todos aquellos números $\beta \in \mathbb{R}$ para los que exista una sucesión $a_1, ..., a_n$ de números reales tales que $a_1^2 \in \mathbb{Q}, a_2^2 \in \mathbb{Q}(a_1), ..., a_n^2 \in \mathbb{Q}(a_1, ..., a_{n-1})$ con $\beta \in \mathbb{Q}(a_1, ..., a_n)$.

Teorema 3.68 *Una longitud $\beta \in \mathbb{R}$ es constructible si y sólo si existe una sucesión $a_1, ..., a_n \in \mathbb{R}$ con $a_1^2 \in \mathbb{Q}, a_2^2 \in \mathbb{Q}(a_1), ..., a_n^2 \in \mathbb{Q}(a_1, ..., a_{n-1})$ tal que $\beta \in \mathbb{Q}(a_1, ..., a_n)$.*

Demostración: Ya tenemos demostrada la implicación $\Leftarrow$. Para ver el recíproco, adaptemos un sistema cartesiano de coordenadas al plano $\mathbb{R}^2$ de manera que la longitud unitaria quede dada por los puntos $(0, 0)$ y $(1, 0)$.

Digamos ahora que un punto (a, b) es constructible si y sólo si sus coordenadas lo son. Esto es razonable, pues con regla y compás se puede trazar una perpendicular a una recta dada desde un punto también dado, en la recta o fuera de ella.

El procedimiento para obtener nuevos puntos a partir de los ya construidos, consiste en intersectar rectas que pasen por puntos construidos, intersectar circunferencias con centros y radios construidos entre sí o bien con rectas construidas. Algebraicamente, esto equivale a resolver sistemas de dos ecuaciones, donde cada ecuación es de una de las formas

$$Ax + By + C = 0 \qquad (1)$$

$$x^2 + y^2 + ax + by + c = 0 \qquad (2)$$

Cuando se resuelve un sistema de dos ecuaciones de tipo (1), se hace dentro del campo de los coeficientes del sistema. Es fácil ver que un sistema de dos ecuaciones de tipo (2) es equivalente a otro sistema con una ecuación de cada tipo.

Si en la ecuación lineal, $B \neq 0$, despejamos y para usarlo en la ecuación cuadrática, obtenemos

$$x^2 + \left(\frac{C}{B} + \frac{A}{B}\ x\right)^2 + ax - b\left(\frac{C}{B} + \frac{A}{B}\ x\right) + c = 0.$$

De manera que de existir soluciones reales, éstas pertenecerán al campo de los coeficientes o a una extensión cuadrática del mismo. □

Corolario 3.69 *Un número real α es constructible si y sólo si $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ es una potencia de dos.*

Identificamos geométricamente a $\mathbb{R}^2$ con $\mathbb{C}$ de la manera usual y consideramos a $\beta = a + bi \in \mathbb{C}$ con $a, b \in \mathbb{R}$. Decimos que β es constructible cuando a y b lo son.

Teorema 3.70 *Las siguientes condiciones en $\beta \in \mathbb{C}$ son equivalentes:*

a) β es constructible.

b) Existe una sucesión $a_1, ..., a_n \in \mathbb{C}$ *tal que*

$$a_1^2 \in \mathbb{Q}, a_2^2 \in \mathbb{Q}(a_1), ..., a_n^2 \in \mathbb{Q}(a_1, ..., a_{n-1}) \text{ con } \beta \in \mathbb{Q}(a_1, ..., a_n).$$

c) β *es algebraico; y la cerradura normal de* $\mathbb{Q}(\beta)$ *tiene grado sobre* $\mathbb{Q}$ *que es una potencia de dos.*

Demostración: *a)* $\Rightarrow$ *b)* Si $\beta = a + bi$ es constructible, con $a, b \in \mathbb{R}$, es porque existen sucesiones $a_1, ..., a_r$ y $b_1, ..., b_s$ como en el enunciado con

$$a \in \mathbb{Q}(a_1, ..., a_r) \quad \text{y} \quad b \in \mathbb{Q}(b_1, ..., b_s);$$

pero entonces la sucesión $a_1, ..., a_r, b_1, ..., b_s, i$ también es como en el enunciado; y satisface

$$\beta \in \mathbb{Q}(a_1, ..., a_r, b_1, ..., b_s, i).$$

b) $\Rightarrow$ *c)* Sean F la cerradura normal de $\mathbb{Q}(\beta)/\mathbb{Q}$ y $G = \text{Gal}(F/\mathbb{Q})$. Para cada $\sigma \in G$, existe una sucesión $\sigma(a_1), ..., \sigma(a_r)$ como en el enunciado, con $\sigma(\beta) \in \mathbb{Q}(\sigma(a_1), ..., \sigma(a_r))$. Aquí vemos que

$$\sigma_1(a_1), ..., \sigma_1(a_r), \sigma_2(a_1), ..., \sigma_2(a_r), ..., \sigma_n(a_1), ..., \sigma_n(a_r)$$

es como en el enunciado con β y todos sus conjugados contenidos en $\mathbb{Q}(\sigma_i(a_j))_{i,j}$. Esto implica que β es algebraico y que $[F : \mathbb{Q}]$ es una potencia de dos.

c) $\Rightarrow$ *a)* Si la cerradura normal F de $\mathbb{Q}(\beta)/\mathbb{Q}$ es de grado 2^t, entonces $G = \text{Gal}(F/\mathbb{Q})$ es un 2-grupo, por tanto nilpotente; y existe una cadena finita de subgrupos de G así:

$$G = G_0 \supset G_1 \supset \cdots \supset G_t = \{1\},$$

donde $G_{i+1} \lhd G_i$ para todo i, con cada cociente G_i/G_{i+1} de orden dos.

Por el Teorema Fundamental de la Teoría de Galois, existe una cadena de campos correspondiente:

$$\mathbb{Q} = K_0 \subset K_1 \subset \cdots \subset K_t = F,$$

con $[K_{i+1} : K_i] = 2$ para todo i.

Como $\beta \in F$, para saber que β es constructible, es suficiente ver por inducción en i, que todo elemento de K_i lo es. Esto es cierto para $K_0 = \mathbb{Q}$. Si $K_{i+1} = K_i(\eta)$, con $\eta = b_1 + b_2 i, \eta^2 = c_1 + c_2 i \in K_i$, donde $b_1, b_2, c_1, c_2 \in \mathbb{R}$, entonces suponiendo que c_1 y c_2 son constructibles, se tiene que b_1 y b_2 satisfacen

$$b_1^2 = \frac{\sqrt{c_1^2 + c_2^2} + c_1}{2} \text{ y } b_2^2 = \frac{\sqrt{c_1^2 + c_2^2} - c_1}{2}.$$

Así, es claro que η es constructible, como lo es todo elemento de K_{i+1}. $\square$

La duplicación del cubo. Aquí el problema es construir un cubo de volumen 2, es decir, la arista de tal cubo. Esto es imposible porque $X^3 - 2$ es irreducible en $\mathbb{Q}[X]$.

Trisección de ángulos Dado un ángulo θ, ¿es posible construir $\theta/3$? La respuesta es negativa, por ejemplo en el caso de $\theta = 60°$. En efecto, $\cos 60° = 1/2$ y $\operatorname{sen} 60° = \sqrt{3}/2$, por lo que un ángulo de $60°$ es constructible. Sin embargo, la constructibilidad de un ángulo de $20°$ es equivalente a la de $\cos 20°$.

La identidad trigonométrica $\cos 3\varphi = 4\cos^3 \varphi - 3\cos\varphi$, produce para $\varphi = 20°$ y $\cos\varphi = \alpha$, la igualdad $4\alpha^3 - 3\alpha - 1/2 = 0$; pero el polinomio $4X^3 - 3X - 1/2$ es irreducible en $\mathbb{Q}[X]$, como se ve al substituir $X = Y/2$ y obtener $(1/2)(Y^3 - 3Y - 1)$. Así, $\cos 20°$ es de grado tres sobre $\mathbb{Q}$; y no es constructible.

La cuadratura del círculo. ¿Es posible construir un círculo de área uno? Como el área de un círculo de radio r es πr^2, la respuesta es negativa porque π es trascendente.

Polígonos regulares constructibles. Un polígono regular de n lados es constructible si sólo si lo es una raíz n-ésima primitiva de la unidad. El siguiente resultado nos da la respuesta a este problema:

Teorema 3.71 *Una raíz n-ésima primitiva de la unidad ζ es constructible si sólo si $n = 2^r p_1 \cdots p_s$ con $r \in \mathbb{N}$ y con $p_1, ..., p_s$ primos de la forma $2^t + 1$ distintos entre sí.*

Demostración: Para $n = 2^r p_1^{r_1} \cdots p_s^{r_s}$ con $p_1, ..., p_s$ primos impares distintos y $r \geq 0; r_1, ..., r_s \geq 1$, el Teorema 3.54 *c*) afirma que el grado de ζ sobre $\mathbb{Q}$ es

$$\varphi(n) = \begin{cases} 2^{r-1}(p_1 - 1)p_1^{r_1 - 1} \cdots (p_s - 1)p_s^{r_s - 1}, \text{ si } r \geq 1 \\ (p_1 - 1)p_1^{r_1 - 1} \cdots (p_s - 1)p_s^{r_s - 1}, \text{ si } r = 0 \end{cases}$$

Como $\mathbb{Q}(\zeta)/\mathbb{Q}$ es normal y $\varphi(n)$ es una potencia de 2 exactamente cuando n es como en el enunciado, obtenemos nuestra conclusión. □

Los números primos de la forma $2^t + 1$ se llaman **primos de Fermat**. Como ejemplos tenemos a $3, 5, 17, 257, 65537$. No se sabe si el número de estos primos es finito.

Ejercicios

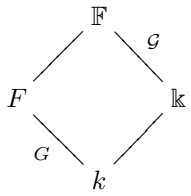
1. Demuestre que $\operatorname{Polmin}(\cos(2\pi/5), \mathbb{Q}) = X^2 + \frac{1}{2}X - \frac{1}{4} = 0$.
2. Demuestre que el polinomio mínimo de $\cos(2\pi/17)$ sobre $\mathbb{Q}$ es

$$X^8 + \frac{1}{2}X^7 - \frac{7}{4}X^6 - \frac{3}{4}X^5 + \frac{15}{16}X^4 + \frac{5}{16}X^3 - \frac{5}{32}X^2 - \frac{1}{32}X + \frac{1}{256} = 0.$$

3.12 Grupos de Galois sobre $\mathbb{Q}$

Sean k un campo, $f(T)$ un polinomio separable de grado n sobre k, F su campo de descomposición y $G = \mathrm{Gal}(F/k)$. Consideramos un conjunto de variables $X_1, ..., X_n$; y los campos de funciones racionales $\mathbb{F} = F(X_1, ..., X_n)$ y $\Bbbk = k(X_1, ..., X_n)$.

Claramente, F/k es una extensión finita de Galois. Como normalidad y separabilidad se preservan ante translación, $\mathbb{F}/\Bbbk$ también es finita de Galois; y $\mathbb{F}$ es campo de descomposición de $f(X)$ sobre $\Bbbk$. Sea $\mathcal{G} = \mathrm{Gal}(\mathbb{F}/\Bbbk)$.



Proposición 3.72 *El morfismo de grupos $\psi : \mathcal{G} \to G$ dado por $\psi(\sigma) = \sigma|_F$ es un isomorfismo.*

Demostración: Como $\sigma|_k = 1_k$, tenemos que σ envía cada raíz de $f(T)$ a otra raíz de $f(T)$. Así, ψ está bien definido.

La biyectividad de ψ es consecuencia de que todo k-automorfismo de F se extiende de manera única a un $\Bbbk$-automorfismo de $\mathbb{F}$. □

Supongamos que $f(T)$ tiene raíces $r_1, ..., r_n$; de manera que tengamos morfismos $G \hookrightarrow S_n$ y $\mathcal{G} \hookrightarrow S_n$. Para $\sigma \in S_n$, definimos

$$u_\sigma = \sum_{i=1}^{n} r_{\sigma(i)} X_i = \sum_{i=1}^{n} r_i X_{\sigma^{-1}(i)}.$$

Observación. Si $\sigma \neq \tau$, entonces $u_\sigma \neq u_\tau$.

Proposición 3.73 *a)* $\mathbb{F} = \Bbbk(u_\sigma)$*, para cualquier* $\sigma \in S_n$*.*

b) $\mathrm{Polmin}(u_\sigma, \Bbbk) = g_\sigma(T)$*, donde*

$$g_\sigma(T) = \prod_{\tau \in \mathcal{G}} (T - \sum_{i=1}^{n} r_{\tau\sigma(i)} X_i).$$

Demostración: Aquí, $\tau g_\sigma(T) = g_\sigma(T)$, para todos $\tau \in \mathcal{G}$, $\sigma \in S_n$, por lo que $g_\sigma(T) \in \Bbbk[T]$ para todo $\sigma \in S_n$. Además, $g_\sigma(T)$ tiene como raíz a u_σ; y es de grado $\circ(\mathcal{G})$. Para obtener *a)* y *b)*, es suficiente saber que todo u_σ es de grado $\geq \circ(\mathcal{G})$ sobre $\Bbbk$; pero esto es consecuencia inmediata de que la órbita de u_σ ante la acción de $\mathcal{G}$ tiene $\circ(\mathcal{G})$ elementos, gracias a la última observación. □

Proposición 3.74 $\mathcal{G} = \{\tau \in S_n \mid \tau g_\sigma(T) = g_\sigma(T) \text{ para todo } \sigma \in S_n\}$.

Demostración: Si escribimos

$$g(T) = \prod_{\sigma \in S_n} (T - u_\sigma) = \prod_{\sigma \in S_n/\mathcal{G}} g_\sigma(T),$$

entonces tendremos que $g_\sigma(T) \in k[X_1, ..., X_n, T]$, para todo $\sigma \in S_n$; y cada $g_\sigma(T)$ es irreducible en $k[X_1, ..., X_n, T]$.

El grupo S_n actúa de manera natural en $k[X_1, ..., X_n, T]$ permutando las variables X_i. Esta da origen a una colección de $k[T]$-automorfismos. Ante esta acción, el subgrupo $\mathcal{G}$ fija a cada factor $g_\sigma(T)$.

Recíprocamente, si $\tau \in S_n$ fija a $g_1(T)$, entonces $\tau \in \mathcal{G}$. □

Sean ahora $k = \mathbb{Q}$ y $f(T) \in \mathbb{Z}[T]$ mónico y separable. Estamos en posición de obtener información valiosa acerca del grupo de Galois de $f(T)$, usando los resultados anteriores, al reducir al polinomio $(\mathrm{mod}\, p)$ para primos p convenientes: Extendemos el morfismo natural $\eta : \mathbb{Z} \to \mathbb{F}_p$ a los anillos de polinomios $\eta' : \mathbb{Z}[X_1, ..., X_n, T] \to \mathbb{F}_p[X_1, ..., X_n, T]$; y escribimos $\overline{f}(T)$ en lugar de $\eta' f(T)$.

Proposición 3.75 *Si $\overline{f}(T)$ es separable y tiene raíces $z_1, ..., z_n$, entonces*

$$g(T) \in \mathbb{Z}[X_1, ..., X_n, T] \quad y \quad \overline{g}(T) = \prod_{\sigma \in S_n} (T - \sum_{i=1}^{n} z_{\sigma(i)} X_i).$$

Demostración: Aquí, $g(T) \in \mathbb{Z}[X_1, ..., X_n, T]$ por el Teorema 2.60. Como $\overline{f}(T)$ es separable, podemos extender η' a $\mathbb{Z}[r_1, ..., r_n, X_1, ..., X_n, T]$; y así obtener la segunda expresión, definiendo $\eta'(r_i) = z_i$ para toda i. □

Teorema 3.76 (Dedekind) *Si $f(T) \in \mathbb{Z}[T]$ es mónico y $\overline{f}(T) \in \mathbb{F}_p[T]$ es separable, entonces*

$$\mathrm{Gal}(\overline{f}(T)/\mathbb{F}_p) < \mathcal{G} = \mathrm{Gal}(f(T)/\mathbb{Q}).$$

Demostración: Como $\mathcal{G}$ es el estabilizador de los factores irreducibles de $g(T)$ en el anillo $\mathbb{Z}[X_1, ..., X_n, T]$, mientras que el grupo $\mathrm{Gal}(\overline{f}(T)/\mathbb{F}_p)$ lo es para los factores irreducibles de $\overline{g}(T)$ en $\mathbb{F}_p[X_1, ..., X_n, T]$, la conclusión es clara. □

Ejemplo. El Teorema 3.50 *c)* nos permite calcular el grupo de Galois G de $f(X) = X^4 + 2X^2 + X + 3$ sobre $\mathbb{Q}$:

$f(X) \equiv X^4 + X + 1 \ (\mathrm{mod}\, 2)$, que es irreducible, pues no tiene raíces y $X^4 + X + 1 \neq (X^2 + X + 1)^2$, sabiendo que $X^2 + X + 1$ es el único polinomio cuadrático irreducible $(\mathrm{mod}\, 2)$. Así, G contiene un 4-ciclo.

$f(X) \equiv X(X^3 + 2X + 1) \ (\mathrm{mod}\, 3)$, con $X^3 + 2X + 1$ irreducible $(\mathrm{mod}\, 3)$. Así, G contiene un 3-ciclo.

Tenemos que G es un grupo tal que contiene un 3-ciclo y un 4-ciclo, además $12 \mid \circ(G)$. Entonces G es un subgrupo normal de S_4 de orden 12 ó 24; al contener un 3-ciclo, los contiene a todos y a A_4. Al contener a un 4-ciclo, que es impar, $G \neq A_4$ implica que $G = S_4$.

A continuación consideraremos el problema inverso de la teoría de Galois: dado un grupo G, exhibir una extensión de Galois de $\mathbb{Q}$, o bien un polinomio en $\mathbb{Q}[X]$ cuyo grupo de Galois sea G. Obtendremos soluciones para los casos del grupo simétrico S_n y de grupos abelianos finitos arbitrarios.

Lema 3.77 *Sea G un subgrupo transitivo de S_n que contiene una transposición y un $(n-1)$-ciclo, entonces $G = S_n$.*

Demostración: Digamos que $\alpha = (23\cdots n) \in G$; y que G también contiene la transposición (ab). Sabiendo que G es transitivo, existe $\sigma \in G$ tal que $\sigma(a) = 1$; pero entonces también existe m tal que $1 < m \leq n$ con $\sigma(ab)\sigma^{-1} = (1m)$.

Al conjugar $(1m)$ con las distintas potencias de α, tenemos que $(1m) \in G$ para todo m con $1 < m \leq n$; pero $\langle\ (1m) \mid 1 < m \leq n\rangle = S_n$, por la Proposición 1.36. Concluimos que $G = S_n$. □

Teorema 3.78 *Para todo entero positivo n, existe un polinomio separable $f(X) \in \mathbb{Q}[X]$, de grado n, tal que $G = \mathrm{Gal}(f(X)/\mathbb{Q}) = S_n$.*

Demostración: Podemos suponer que $n \geq 3$. Gracias a los Teoremas 3.48 y 3.36 que afirman que existen extensiones de $\mathbb{F}_p$ de cualquier grado para cualquier primo p; y que estas son simples, puede verse que existen $f_1(X) \in \mathbb{F}_2[X]$ mónico, irreducible de grado n, $g_2(X) \in \mathbb{F}_3[X]$ mónico, irreducible de grado $n-1$; así como $f_3(X) \in \mathbb{F}_5[X]$ mónico, de grado n, con factorización irreducible consistente en un polinomio cuadrático, junto con uno o dos factores de grado impar. Definimos $f_2(X) = Xg_2(X)$ y observamos que es posible tener todo $f_i(X)$ separable.

Por el Teorema Chino del Resíduo, existe $f(X) \in \mathbb{Z}[X]$ tal que

$$f(X) \equiv f_1(X) \pmod 2, f(X) \equiv f_2(X) \pmod 3, f(X) \equiv f_3(X) \pmod 5.$$

El Teorema 3.76 garantiza que G contiene un n-ciclo, un $(n-1)$-ciclo y una permutación β con descomposición cíclica $(2, n-2)$ ó bien $(2, a, b)$, con $n-2$ impar ó bien a y b impares. De esta manera, β^{n-2} ó bien β^{ab} es una transposición en G. El lema implica que $G = S_n$. □

Lema 3.79 *Sean p, n, a enteros con p primo, $p \nmid n$ y tales que $p \mid \Phi_n(a)$, donde $\Phi_n(X)$ es el n-ésimo polinomio ciclotómico. Entonces $p \equiv 1 \pmod n$.*

Demostración: Aquí, a es una raíz n-ésima primitiva de la unidad sobre $\mathbb{F}_p$. Como $a^{p-1} = 1$, es inmediato que $n \mid (p-1)$. □

El siguiente resultado es un caso sencillo de un teorema de Dirichlet.

Proposición 3.80 *Sea n un entero positivo. Existe un número infinito de primos de la forma $mn+1$, con $m \in \mathbb{N}$.*

Demostración: El término constante de $\Phi_n(X) \in \mathbb{Z}[X]$ tiene valor absoluto que es el de la norma de una raíz de la unidad, por lo que se tiene $\Phi_n(X) = X^{\varphi(n)} + \cdots + (\pm 1) \in \mathbb{Z}[X]$. Supongamos que $p_1, ..., p_r$ es la lista completa de primos $p \equiv 1 \pmod{n}$. Para todo entero positivo i, tenemos

$$\Phi_n(n^i p_1 \cdots p_r) \equiv \pm 1 \pmod{n, p_1, ..., p_r}.$$

Escogiendo i adecuadamente, podemos asegurar que

$$m = \Phi_n(n^i p_1 \cdots p_r) \neq \pm 1.$$

Sea p un primo que divida a m, entonces $p \nmid n$. Por el lema, $p \equiv 1 \pmod{n}$, en contradicción con $p \neq p_j$, para todo j. □

Teorema 3.81 *Todo grupo abeliano finito G es el grupo de Galois sobre $\mathbb{Q}$ de una extensión subciclotómica.*

Demostración: G es el producto directo de grupos cíclicos de órdenes $n_1, ..., n_r$. Sean $p_1, ..., p_r$ primos distintos tales que $p_i \equiv 1 \pmod{n_i}$, para todo i. Escribamos $m = p_1 \cdots p_r$.

Sea ζ una raíz m-ésima primitiva de la unidad sobre $\mathbb{Q}$. Entonces $\mathbb{Q}(\zeta)/\mathbb{Q}$ es una extensión finita de Galois con grupo isomorfo a $Z_{p_1-1} \times \cdots \times Z_{p_r-1}$. Existen subgrupos $H_i < Z_{p_i-1}$ de índice n_i, para todo i.

Sea $H = H_1 \times \cdots \times H_r < \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$. Tenemos que $\mathbb{Q}(\zeta)^H/\mathbb{Q}$ es una extensión finita de Galois de $\mathbb{Q}$, cuyo grupo es isomorfo con G. □

Ejercicio

1. Encuentre un polinomio cuyo grupo de Galois sobre $\mathbb{Q}$ sea S_5.

3.13 Ejercicios Generales

1. Sean $p > 2$ un número primo y ζ una raíz p-ésima primitiva de la unidad. Demuestre que existe un único campo E tal que

$$\mathbb{Q} \subset E \subseteq \mathbb{Q}(\zeta) \quad \text{y} \quad [E : \mathbb{Q}] = 2.$$

Además, $E \subseteq \mathbb{R} \Leftrightarrow p \equiv 1 \pmod{4}$.

2. Sea p número primo impar.

   a) Si ω es una raíz p-ésima primitiva de la unidad, demuestre que

$$\prod_{i=1}^{p-1}(1-\omega^i)=p.$$

   b) Demuestre que el discriminante de X^p-1 es $(-1)^{(p-1)/2}p^p$.

   c) Demuestre que el discriminante de $\Phi_p(X)$ es $(-1)^{(p-1)/2}p^{p-2}$.

3. Demuestre que el discriminante de X^n-1 es $(-1)^{(n-1)/2}n^n$, si n es impar; o bien $(-1)^{(n/2)-1}n^n$, si n es par.

4. Sean p número primo impar y $1\leq r\in\mathbb{N}$.

   a) Demuestre que

$$\Phi_{p^r}(X)=X^{p^r-p^{r-1}}+X^{p^r-2p^{r-1}}+\cdots+X^{p^{r-1}}+1.$$

   b) Demuestre que

$$\prod_{\substack{\omega\ \text{primitiva}\\ \omega^{p^r}=1}}(1-\omega)=p.$$

   c) Demuestre que si ζ es raíz de $X^{p^{r-1}}-1$, entonces

$$\prod_{\substack{\omega\ \text{primitiva}\\ \omega^{p^r}=1}}(\zeta-\omega)=p.$$

   d) Demuestre que

$$\prod_{\substack{\zeta^{p^{r-1}}=1\\ \omega^{p^r}=1,\ \omega\ \text{primitiva}}}(\zeta-\omega)^2=p^{2p^{r-1}}.$$

   e) Demuestre que el discriminante de $\Phi_{p^r}(X)$ es

$$(-1)^{(p-1)/2}p^{(p^{r-1})(pr-r-1)}.$$

5. Escribiendo $d_r=\operatorname{discr}(X^r-1)$, demuestre que

$$\operatorname{discr}(\Phi_n(X))=\prod_{r|n}d_r^{\mu(n/r)}\prod_{\substack{r|n\\ r\neq 1}}\left(\frac{n}{r}\right)^{2r\mu(n/r)}.$$

Capítulo 4
Algebra Lineal

4.1 Módulos Libres

En esta sección suponemos que k es un anillo conmutativo con 1 y que M es un k-módulo izquierdo. En estas condiciones, tenemos que:

1. Toda intersección de submódulos de M es otro submódulo.
2. Dado un subconjunto A de M, el submódulo generado por A, escrito (A), es la intersección de los submódulos de M que contienen al conjunto A.
3. Si $M_1, ..., M_r$ son submódulos de M, entonces $(M_1 \cup \cdots \cup M_r) = M_1 + \cdots + M_r$.

Se dice que M es finitamente generado cuando admite a un conjunto finito como generador. Se dice que M es **suma directa** de sus submódulos $M_1, ..., M_r$ cuando $M_1 + \cdots + M_r = M$ y también $(M_1 + \cdots + M_{i-1}) \cap M_i = (0)$ para todo $1 < i \leq r$. Esto se escribe así: $M = M_1 \oplus \cdots \oplus M_r$.

Un **morfismo** de k-módulos $f : M \to N$, o **función lineal**, es una función tal que $f(am+bn) = af(m)+bf(n)$ siempre que $a, b \in k$; $m, n \in M$. Una función $f : M_1 \times \cdots \times M_r \to N$ es **multilineal** cuando todas las funciones $f_i : M_i \to N$ dadas por $f_i(x) = f(m_1, ..., m_{i-1}, x, m_{i+1}, ..., m_r)$ son lineales para todos $m_j \in M_j,\ 1 \leq i, j \leq r,\ j \neq i$.

Se dice que un k-módulo finitamente generado M es **libre** cuando M es la suma directa de (un número finito de) copias de k; si el número de copias de k es n, escribimos $M = k^{(n)}$.

Un conjunto $\{u_1, ..., u_r\} \subseteq M$ es **linealmente independiente** cuando $\sum_{i=1}^{r} a_i u_i = 0 \Rightarrow a_i = 0,$ para todos $a_i \in k,\ 1 \leq i \leq r$.

Un subconjunto $A = \{u_1, ..., u_r\} \subseteq M$ es una **base** de M cuando es linealmente independiente y genera a M. Esto es equivalente a exigir que todo elemento $v \in M$ se pueda expresar de manera única como combinación lineal de A, es decir, que existan $c_i \in k$ únicos tales que $v = \sum_{i=1}^{r} c_i u_i$.

Observación. En el caso en que k es un campo, todo módulo es libre, al ser un espacio vectorial.

Teorema 4.1 *a) Un k-módulo M es libre si y sólo si tiene una base.*

b) Dos bases de un mismo módulo (libre) tienen el mismo número de elementos.

Demostración: *a)* $k^{(n)}$ admite como base al conjunto $\{\epsilon_1, ..., \epsilon_n\}$, donde $\epsilon_1 = (1, 0, ..., 0), \epsilon_2 = (0, 1, 0, ..., 0)$, etc.

Si M admite a $\{u_1, ..., u_n\}$ como base, entonces la función $f : k^{(n)} \to M$ dada por $f(a_1, ..., a_n) = a_1u_1 + \cdots + a_nu_n$ es un isomorfismo de k-módulos.

b) Aquí podemos adaptar una versión del Teorema de Jordan-Hölder, o bien reducimos el problema al caso ya conocido para campos:

Existe un ideal máximo $\mathfrak{m}$ de k, de manera que $F = k/\mathfrak{m}$ es un campo. Escribimos $\overline{u}$ la imagen de $u \in M$ en $M/\mathfrak{m}M$. Dada una base $\{u_1, ..., u_r\}$ de M, tenemos que $B = \{\overline{u}_1, ..., \overline{u}_r\}$ genera a $M/\mathfrak{m}M$ como $k/\mathfrak{m}$-módulo, es decir, como espacio vectorial. Veamos que B es linealmente independiente:

$$\sum_{i=1}^{r} \overline{a}_i\overline{u}_i = 0, \text{ con } \overline{a}_i \in F \text{ y } a_i \in k \Rightarrow \sum_{i=1}^{r} a_iu_i \in \mathfrak{m}M \Rightarrow$$

$a_i \in \mathfrak{m}$ para todo $1 \leq i \leq r$, es decir, $\overline{a}_i = 0$ para todo $1 \leq i \leq r$.

Así, $M/\mathfrak{m}M$ es un espacio vectorial sobre F de dimensión r, de donde se obtiene la unicidad de r. □

Decimos que un módulo libre tiene **rango** n cuando admite una base con n elementos. Acabamos de ver que esto ocurre si y sólo si $M \cong k^{(n)}$.

Observación. El objeto de estudio del Algebra Lineal son los módulos libres y sus morfismos, comunmente restringidos al caso en que k es un campo. El rango de un módulo libre sobre un campo es su dimensión como espacio vectorial.

Sea $f : M \to N$ un morfismo de módulos libres con rango $M = m$ y rango $N = n$. Elegimos una base $\{u_1, ..., u_m\}$ de M y una base $\{v_1, ..., v_n\}$ de N. Esta elección de bases nos permite asociarle una matriz A de tamaño $m \times n$ al morfismo f así:

$$A = (a_{ij}), \text{ donde } f(u_i) = \sum_{j=1}^{n} a_{ij}v_j, \text{ para } 1 \leq i \leq m.$$

En estas condiciones, si $x = \sum_{i=1}^{m} x_iu_i$ con $y = f(x) = \sum_{j=1}^{n} y_iv_i$; de manera que $x \in M$ tiene coordenadas $x_1, ..., x_m$, mientras que las de $y \in N$ son $y_1, ..., y_n$, se cumple la identidad matricial $(x_1, ..., x_m)A = (y_1, ..., y_n)$, como puede verificarse en los casos $x = u_i$.

Obtenemos el morfismo $\Psi : \mathrm{Hom}_k(M, N) \to M_{m\times n}(k)$, donde $M_{m\times n}(k)$ es el conjunto de matrices $m \times n$ con coeficientes en k y $\mathrm{Hom}_k(M, N)$ es el k-módulo de funciones lineales $M \to N$.

Aquí, $\Psi(f) = A$ en la situación anterior; y vemos que Ψ es un isomorfismo de k-módulos, pues se puede obtener toda matriz $A \in M_{m\times n}(k)$ como imagen de algún morfismo $f : M \to N$, mientras que $\ker \Psi = (0)$.

Cuando $M = N$, elegimos solamente una base de M, entonces resulta que $\Psi : \mathrm{End}_k(M) = \mathrm{Hom}_k(M, M) \to M_n(k)$ es un antiisomorfismo de anillos:

$$\Psi(g \circ f) = \Psi(f)\Psi(g), \text{ para } f, g \in \underset{k}{\mathrm{End}}(M).$$

Proposición 4.2 *Si $f \in \mathrm{End}_k(M)$ es tal que envía una base de M a otra base de M, entonces f es un automorfismo de M.*

Demostración: Sean $\{u_1, ..., u_n\}$ y $\{v_1, ..., v_n\}$ dos bases de M tales que $f(u_i) = v_i$ para todo $1 \le i \le n$, entonces $\mathrm{Im}\, f = (v_1, ..., v_n) = M$, mientras que $\ker f = \{\sum_{i=1}^n a_i u_i \mid f(\sum_{i=1}^n a_i u_i) = \sum_{i=1}^n a_i v_i = 0\} = (0)$. Así, f es biyectivo. Es fácil ver que f^{-1} es lineal. □

Teorema 4.3 *Si al morfismo de k-módulos libres $f : M \to N$, donde $\mathrm{rango}\, M = m$ y $\mathrm{rango}\, N = n$, le corresponde la matriz $A = (a_{ij})$ con respecto a las bases $\{u_1, ..., u_m\}$ de M y $\{v_1, ..., v_n\}$ de N, entonces también le corresponde la matriz PAQ^{-1}, con respecto a las bases $\{u'_1, ..., u'_m\}$ de M y $\{v'_1, ..., v'_n\}$ de N dadas por $u'_i = \sum_{j=1}^m p_{ij} u_j$ y $v'_r = \sum_{r=1}^n q_{rs} v_s$, donde $P = (p_{ij})$ es de tamaño $m \times m$ y $Q = (q_{rs})$ es de tamaño $n \times n$.*

Demostración: Como las matrices P y Q están asociadas a cambios de base, el resultado anterior garantiza que son invertibles. Así podemos escribir $Q^{-1} = (t_{rs})$; y también $v_r = \sum_{s=1}^n t_{rs} v'_s$, para $1 \le r \le n$. La demostración concluye con el cálculo siguiente:

$$f(u'_i) = f(\sum_{j=1}^m p_{ij} u_j) = \sum_{j=1}^m p_{ij} f(u_j) = \sum_{j,r} p_{ij} a_{jr} v_r = \sum_{j,r,s} p_{ij} a_{jr} t_{rs} v'_s. \quad \square$$

Se dice que dos matrices $A, B \in M_{m\times n}(k)$ son **equivalentes** cuando existen matrices invertibles $P \in M_m(k)$ y $Q \in M_n(k)$ tales que $B = PAQ$. Acabamos de ver que esto sucede exactamente cuando A y B representan a un mismo morfismo de k-módulos libres $k^{(m)} \to k^{(n)}$, donde los cambios de base están dados por P y Q^{-1} respectivamente.

Dualidad

Cuando V es un k-módulo libre y k es un anillo conmutativo, el k-módulo $\mathrm{Hom}_k(V, k)$ es el **módulo dual** de V, escrito V^*.

Teorema 4.4 *a) Si V es un k-módulo libre con base $\{\epsilon_1, ..., \epsilon_n\}$, entonces V^* admite como base a $\{\epsilon_1^*, ..., \epsilon_n^*\}$, donde $\epsilon_i^*(\epsilon_j) = \delta_{ij}$ para $1 \le i, j \le n$.*

b) La función $h : V \to V^{**}$ dada por $h(u)(v^*) = v^*(u) \in k$ para todos $u \in V, v^* \in V^*$ es un isomorfismo de k-módulos.

c) Sea $T : V \to V$ lineal tal que $T(\epsilon_i) = \sum_{j=1}^{n} a_{ij}\epsilon_j$, para todo $1 \leq i \leq n$. Definimos $T^* : V^* \to V^*$ lineal así: $T^*(f)(v) = f(T(v))$, para $f \in V^*$ y $v \in V$. Entonces la matriz asociada a T^* con respecto a la base $\{\epsilon_1^*, ..., \epsilon_n^*\}$ de V^* es A^t, la transpuesta de $A = (a_{ij})$.

Demostración: *a)* Sea $f \in V^*$ tal que $f(\epsilon_i) = a_i \in k$ para $1 \leq i \leq n$, entonces $(f - \sum_{i=1}^{n} a_i\epsilon_i^*)(\epsilon_j) = 0$, para todo $1 \leq j \leq n$. Así, tenemos que $f = \sum_{i=1}^{n} a_i\epsilon_i^*$, por lo que $(\epsilon_1^*, ..., \epsilon_n^*) = V^*$.

Por otra parte, $\sum_{i=1}^{n} c_i\epsilon_i^* = 0 \Rightarrow c_j = \sum_{i=1}^{n} c_i\epsilon_i^*(\epsilon_j) = 0$ para todo $1 \leq j \leq n$. Así, $\{\epsilon_1^*, ..., \epsilon_n^*\}$ es linealmente independiente.

b) h es claramente lineal: $h(au_1 + bu_2)(v^*) = v^*(au_1 + bu_2) = av^*(u_1) + bv^*(u_2) = (ah(u_1) + bh(u_1))(v^*)$, para todos $a, b \in k, u_1, u_2 \in V, v^* \in V^*$.

Ahora bien, $u = \sum_{i=1}^{n} c_i\epsilon_i \in \ker h \Rightarrow 0 = h(u)(\epsilon_j^*) = \epsilon_j^*(\sum_{i=1}^{n} c_i\epsilon_i) = c_j$, para todo $1 \leq j \leq n$. Así, $u = 0$ y h es inyectiva.

Finalmente, $(\epsilon_1^{**}, ..., \epsilon_n^{**}) = V^{**}$ y $h(\epsilon_i) = \epsilon_i^{**}$ para todo $1 \leq i \leq n$. Así, h es suprayectiva.

c) Sea $B = (b_{ij})$ la matriz tal que $T^*(\epsilon_i^*) = \sum_{j=1}^{n} b_{ij}\epsilon_j^*$, calculamos b_{ir} para $1 \leq i, r \leq n$:

$$b_{ir} = \sum_{j=1}^{n}(b_{ij}\epsilon_j^*)(\epsilon_r) = (T^*\epsilon_i^*)(\epsilon_r) = \epsilon_i^*(T(\epsilon_r)) = \epsilon_i^*(\sum_{j=1}^{n} a_{rj}\epsilon_j) = a_{ri}. \square$$

Ejercicios

1. Sea k un anillo conmutativo con 1. Demuestre que $\mathrm{Hom}_k(k^{(m)}, k^{(n)})$ es un módulo libre de rango mn.

2. Si $M = M_1 + \cdots + M_r$. Demuestre que $M = M_1 \oplus \cdots \oplus M_r$ si y sólo si $(M_1 + \cdots + M_{i-1} + M_{i+1} + \cdots + M_r) \cap M_i = (0)$ para todo $1 \leq i \leq r$.

3. Exhiba un conjunto $\{v_1, ..., v_n\} \subseteq k^{(n)}$ linealmente independiente que no sea base de $k^{(n)}$, para k un anillo conmutativo que no sea campo.

4. Sean V y W espacios vectoriales de dimensión finita. Demuestre que $\dim V + \dim W = \dim(V + W) + \dim(V \cap W)$.

5. Sea $T : M \to L$ lineal. Defina una **transformación dual** adecuada $T^* : L^* \to M^*$ que sea lineal; y demuestre que

   a) T es inyectivo si y sólo si T^* es suprayectivo.

   b) $T(v) = w$ tiene al menos una solución v para cada $w \in L$ si y sólo si $T^*(w^*) = v^*$ tiene cuando más una solución w^* para cada $v^* \in M^*$.

   c) Cuando $M = L$ y $T, S \in \mathrm{End}(M)$, entonces $(ST)^* = T^*S^*$ y $1^* = 1$.

   d) Cuando $M = L$ y T es invertible, entonces T^* es invertible y $(T^*)^{-1} = (T^{-1})^*$.

4.2 Algebras

En esta sección, suponemos que k es un anillo conmutativo con uno. Un **álgebra sobre** k, también llamado **k-álgebra** es un morfismo de anillos $f : k \to A$ tal que $\operatorname{Im} f \subseteq Z(A)$. Cuando el morfismo f está entendido, abusivamente se dice que el álgebra es el anillo A, que en todo caso posee una estructura adicional de k-módulo.

Ejemplos. Algunos de los objetos más importantes que se estudian en Algebra son precisamente k-álgebras:

1. Si F/k es una extensión de campos, entonces la inclusión $k \hookrightarrow F$ es un k-álgebra.

2. El anillo de polinomios $A = k[X_1, ..., X_n]$ en n variables, donde el morfismo estructural $f : k \to A$ envía los elementos de k a los polinomios constantes.

3. El anillo $\operatorname{End} V$, donde V es un módulo libre sobre el anillo conmutativo k; y donde $k \to \operatorname{End} V$ envía cada elemento $a \in k$ a la multiplicación izquierda por a.

4. El anillo de matrices cuadradas $M_n(k)$, donde el morfismo $f : k \to A$ envía los elementos de k a las matrices escalares, esto es, $f(c) = (c\delta_{ij})$.

5. Sea G un grupo. El **k-álgebra de grupo** $k[G]$ se construye así:

$$k[G] = \{\sum_{g\in G} c_g g \mid c_g \in k, c_g \neq 0 \text{ en un subconjunto finito de } G\},$$
$$\text{donde } \sum_{g\in G} a_g g + \sum_{g\in G} b_g g = \sum_{g\in G} (a_g + b_g)g \quad \text{y}$$
$$(\sum_{h\in G} a_h h)(\sum_{t\in G} b_t t) = \sum_{g\in G} c_g g, \text{ con } c_g = \sum_{ht=g} a_h b_t.$$

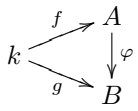
   Aquí, $f : k \to k[G]$ está dado por $f(a) = a1$, para $a \in k$, donde 1 es la identidad del grupo.

6. En el ejemplo anterior, es suficiente que G sea un monoide para producir un álgebra, el **k-álgebra de monoide** $k[G]$. Así, cuando G es el monoide libre abeliano generado por $X_1, ..., X_n$, escrito multiplicativamente, resulta que $k[G]$ es el anillo de polinomios $k[X_1, ..., X_n]$.

7. El álgebra de **cuaternios reales** $\mathbb{R} \to \mathbb{H}$, donde el anillo $\mathbb{H}$ es el del Ejemplo 9 de la Sección 2.1; y donde el morfismo estructural es $\varphi : \mathbb{R} \to \mathbb{H}$ con $\varphi(a) = a + 0i + 0j + 0k$ para todo $a \in \mathbb{R}$.

Un k-álgebra $f : k \to A$ es **asociativo, conmutativo**, etc. según lo sea el anillo A. Todos los ejemplos anteriores son asociativos.

Un morfismo de k-álgebras $f : k \to A$ y $g : k \to B$ es un morfismo de anillos $\varphi : A \to B$ que hace conmutativo al diagrama



Dada una colección posiblemente infinita de conjuntos $\{A_i \mid i \in I\}$, se define el **producto cartesiano** de esta colección: $\prod_{i\in I} A_i$ como el conjunto de funciones $f : I \to \cup_{i\in I} A_i$ tales que $f(i) \in A_i$ para todo $i \in I$. También definimos las **proyecciones** $\pi_i : \prod_{i\in I} A_i \to A_i$ así: $\pi_i(f) = f(i)$.

Observemos que dado un conjunto C y dada una colección de funciones $\{g_i : C \to A_i \mid i \in I\}$, existe una única función $g : C \to \prod_{i\in I} A_i$ tal que $\pi_i \circ g = g_i$ para todo $i \in I$. La función g está dada por $g(c) = f_c \in \prod_{i\in I} A_i$, con $f_c(i) = g_i(c)$ para todo $i \in I$.

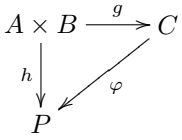
Dada una colección posiblemente infinita de k-módulos $\{M_i \mid i \in I\}$, se define el **producto directo** de esta colección: como el k-módulo con operaciones definidas sobre el producto cartesiano $\prod_{i\in I} M_i$ de los M_i así: $(f + g)(i) = f(i) + g(i)$ y $(cf)(i) = cf(i)$, para $i \in I; c \in k; f, g \in \prod_{i\in I} M_i$. Al k-módulo así obtenido también lo escribimos $\prod_{i\in I} M_i$.

El producto directo de k-módulos también satisface una propiedad universal: Dados un k-módulo C y k-morfismos $\{g_i : C \to M_i \mid i \in I\}$, existe un único k-morfismo $g : C \to \prod_{i\in I} M_i$ tal que $\pi_i \circ g = g_i$ para todo $i \in I$.

Para la misma colección $\{M_i \mid i \in I\}$, se define la **suma directa** $\coprod_{i\in I} M_i$ como el k-submódulo del producto directo, consistente de aquellas funciones $f : I \to \cup_{i\in I} A_i$ tales que $f(i) \neq 0$ para subconjuntos finitos de I. Cuando el conjunto índice I es finito, también usamos (y preferimos) la notación $\bigoplus_{i\in I} M_i$ para la suma directa. Para cada $i \in I$, tenemos la **inyección canónica** $j_i : M_i \to \coprod_{i\in I} M_i$, dada por $j_i(m) = f$, donde $f(i) = m$ y $f(t) = 0$ para $t \neq i$. Este k-morfismo es inyectivo, pues $\pi_i \circ j_i = id_{M_i}$.

Se dice que M es **libre** cuando $M = \coprod_{i\in I} M_i$, con $M_i = k$ para todo $i \in I$, para algún conjunto I. Esta definición generaliza a la dada en la sección anterior para módulos finitamente generados.

Sean A y B dos k-módulos. El **producto tensorial** de A y B es un k-módulo C para el que existe una función bilineal $g : A \times B \to C$ tal que dados un k-módulo P y una función bilineal $h : A \times B \to P$, siempre existe un único k-morfismo (función lineal de k-módulos) $\varphi : C \to P$ tal que el siguiente diagrama conmuta:



A partir de la definición, el producto tensorial es único hasta isomorfismo, en caso de existir, por razones puramente filosóficas: Si C y D son dos objetos adecuados, en este caso k-módulos, que cumplen la condición universal requerida, entonces existen k-morfismos $C \xrightarrow{\varphi} D$ y $D \xrightarrow{\psi} C$, con $\varphi \circ \psi = id_D$ y $\psi \circ \varphi = id_C$, de manera que C y D son isomorfos. A continuación construimos el producto tensorial de los k-módulos A y B.

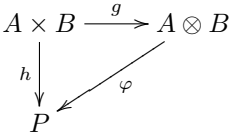
Sea M el k-módulo libre generado por el (conjunto) producto cartesiano $A \times B$ y sea N el submódulo de M generado por los elementos de la forma

$$\begin{aligned} r(m,n) - (rm,n), \quad r(m,n) - (m,rn), \\ (m_1+m_2,n) - (m_1,n) - (m_2,n), \\ (m,n_1+n_2) - (m,n_1) - (m,n_2), \\ \text{para } r \in k; \quad m,m_1,m_2 \in A; \quad n,n_1,n_2 \in B. \end{aligned}$$

Sea $C = M/N$ y sea g la composición de la inyección $A \times B \hookrightarrow M$ con la proyección $M \to M/N = C$. Entonces g es una función bilineal de k-módulos, que claramente cumple con la propiedad universal requerida. En lugar de C, escribimos el producto tensorial de A y B así: $A \otimes_k B$, omitiendo el subídice k cuando quede entendido. Si $a \in A$ y $b \in B$, escribimos $a \otimes b$ en lugar de $g(a,b)$.

Observaciones. De manera análoga, a partir de los k-módulos $A_1, ..., A_n$, es posible construir el producto $A_1 \otimes \cdots \otimes A_n$.

Es muy importante notar que dados k-módulos A, B y P, existe una biyección del conjunto $\text{Bil}(A \times B, P) = \{A \times B \xrightarrow{h} P \mid h \text{ bilineal}\}$ al conjunto $\{A \otimes B \xrightarrow{\varphi} P \mid \varphi \text{ lineal}\} = \text{Hom}(A \otimes B, P)$, donde g es el morfismo usado para construir al producto tensorial y $h = \varphi \circ g$ en el diagrama



Dados k-módulos A, B, en general, no es inmediata la estructura del producto $A \otimes B$. El párrafo anterior nos debe dar una idea de la complejidad del problema, así como de la posible estructura de $A \otimes B$.

Una sucesión de k-módulos y morfismos $\cdots \to N \xrightarrow{f} M \xrightarrow{g} P \to \cdots$ es **exacta** en M cuando $\text{Im}\, f = \ker g$. Tendremos una **sucesión exacta** cuando lo sea en todos sus módulos. Por ejemplo, la exactitud de la sucesión

$$N \xrightarrow{f} M \xrightarrow{g} P \to 0,$$

significa que además de $\text{Im}\, f = \ker g$, se tiene que g es suprayectivo.

A continuación reunimos algunas propiedades del producto tensorial.

Proposición 4.5 *Si M, N, P y Q son k-módulos, entonces*

1. $k \otimes_k M \cong M$.
2. $M \otimes (N \otimes P) \cong M \otimes N \otimes P \cong (M \otimes N) \otimes P$.
3. $M \otimes N \cong N \otimes M$.
4. $M \otimes (N \oplus P) \cong (M \otimes N) \oplus (M \otimes P)$.
5. *Si* $N \xrightarrow{f} M \xrightarrow{g} P \to 0$ *es una sucesión exacta, entonces también lo es* $Q \otimes N \xrightarrow{1\otimes f} Q \otimes M \xrightarrow{1\otimes g} Q \otimes P \to 0$.

Demostración:

1. El morfismo $f : k \times M \to M$, dado por $f(a, m) = am$ es bilineal; y es tal que para cualesquiera k-módulo L y k-morfismo bilineal $h : k \times M \to L$, existe un único morfismo φ que hace conmutativo al diagrama

$$\begin{array}{ccc} k \times M & \xrightarrow{f} & M \\ {\scriptstyle h}\downarrow & \swarrow{\scriptstyle \varphi} & \\ L & & \end{array}$$

2. Veamos que $M \otimes (N \otimes P) \cong M \otimes N \otimes P$. Existe una única función bilineal $g : M \times (N \otimes P) \to M \otimes N \otimes P$ que cumple la condición $g(m, n \otimes p) = m \otimes n \otimes p$, para $m \in M, n \in N, p \in P$. Dado un k-módulo L, es suficiente observar que a cada $h \in \mathrm{Bil}(M \times (N \otimes P), L)$, le corresponde un morfismo único $\varphi \in \mathrm{Hom}(M \otimes N \otimes P, L)$ tal que $h = \varphi \circ g$:

$$\begin{array}{ccc} M \times (N \otimes P) & \xrightarrow{g} & M \otimes N \otimes P \\ {\scriptstyle h}\downarrow & \swarrow{\scriptstyle \varphi} & \\ L & & \end{array}$$

   Esto nos dice que $M \otimes N \otimes P$ es (isomorfo con) el producto tensorial $M \otimes (N \otimes P)$. La otra afirmación admite una demostración análoga.

3. Aquí es suficiente identificar $\mathrm{Bil}(M \times N, L)$ con $\mathrm{Bil}(N \times M, L)$ y con $\mathrm{Hom}(N \otimes M, L)$, observando que la identificación se hace ante $g : M \times N \to N \otimes M$, donde $g(m, n) = n \otimes m$ para $m \in M, n \in N$.

4. Sea $g : M \times (N \oplus P) \to (M \otimes N) \oplus (M \otimes P)$ el k-morfismo dado por $g(m, (n, p)) = (m \otimes n, m \otimes p)$, para $m \in M, n \in N, p \in P$. Dado un k-módulo L, es g la función bilineal usada para identificar

$$\begin{aligned} &\mathrm{Bil}(M \times (N \oplus P), L) \\ \leftrightarrows\ &\mathrm{Bil}(M \times N, L) \oplus \mathrm{Bil}(M \times P, L) \\ \leftrightarrows\ &\mathrm{Hom}(M \otimes N, L) \oplus \mathrm{Hom}(M \otimes P, L). \end{aligned}$$

5. Dado un generador $q \otimes p$ de $Q \otimes P$, existe $m \in M$ tal que $g(m) = p$, de manera que $(1 \otimes g)(q \otimes m) = q \otimes p$, por lo que $1 \otimes g$ es suprayectivo.

   Por otra parte, $g \circ f = 0 \Rightarrow (1 \otimes g) \circ (1 \otimes f) = 0$, de donde obtenemos que $\mathrm{Im}(1 \otimes f) \subseteq \ker(1 \otimes g)$. Esta inclusión nos permite definir un morfismo $\varphi : (Q \otimes M)/(\mathrm{Im}(1 \otimes f)) \to Q \otimes P$ satisfaciendo $\varphi[(q \otimes m) + \mathrm{Im}(1 \otimes f)] = q \otimes g(m)$. Para concluir, veremos que φ es inyectivo:

   Dados $q \in Q$ y $p \in P$, elegimos $m \in M$ tal que $g(m) = p$. Afirmamos que la clase $(q \otimes m) + \mathrm{Im}(1 \otimes f)$ es independiente de la elección de m, pues si $m' \in M$ cumple con $g(m') = p$, entonces $m - m' \in \ker g$, por lo que existe $z \in N$ con $f(z) = m - m'$; y así $(q \otimes m) + \mathrm{Im}(1 \otimes f) = (q \otimes m') + \mathrm{Im}(1 \otimes f)$. Esto nos permite definir un k-morfismo $\eta : Q \times P \to (Q \otimes M)/(\mathrm{Im}(1 \otimes f))$ tal que $\eta(q, p) = q \otimes m + (\mathrm{Im}(1 \otimes f))$, para cualquier $m \in M$ tal que $g(m) = p$; pero entonces existe otro k-morfismo $\psi : Q \otimes P \to (Q \otimes M)/(\mathrm{Im}(1 \otimes f))$ tal que $\psi(q \otimes p) = q \otimes m + (\mathrm{Im}(1 \otimes f))$ con $g(m) = p$. Como $\psi \circ \varphi = 1$, tenemos nuestra conclusión. □

Observaciones. La notación $1 \otimes f$ usada en la proposición, es abusiva, pues no se refiere a ningún producto tensorial de funciones; sino a la función inducida por 1 y f en un producto tensorial de módulos.

El producto tensorial de dos módulos da origen a reestructuraciones. Por ejemplo, $Z_m \otimes_{\mathbb{Z}} Z_n = (0)$, si Z_n es el grupo cíclico de orden n y $(m, n) = 1$, pues cualquier $a \otimes b \in Z_m \otimes Z_n$ es una $\mathbb{Z}$-combinación lineal de $m(a \otimes b) = (ma \otimes b) = 0$ y de $n(a \otimes b) = (a \otimes nb) = 0$.

No es cierto que para todo anillo conmutativo k, toda sucesión exacta $N \xrightarrow{g} M \xrightarrow{f} P$ de k-módulos y todo k-módulo L, siempre se tenga exactitud de la sucesión $L \otimes N \xrightarrow{1 \otimes g} L \otimes M \xrightarrow{1 \otimes f} L \otimes P$. Por ejemplo, la sucesión $0 \to \mathbb{Z} \xrightarrow{f} \mathbb{Z}$ de $\mathbb{Z}$-módulos, con $f(z) = 2z$ para todo $z \in \mathbb{Z}$ es exacta; pero si $L = \mathbb{Z}/2\mathbb{Z}$, entonces

$$0 \to L \otimes \mathbb{Z} \xrightarrow{1 \otimes f} L \otimes \mathbb{Z}$$

no es exacta, pues $(1 \otimes f)(a \otimes b) = a \otimes 2b = 2a \otimes b = 0$, por lo que $1 \otimes f = 0$, mientras que $L \otimes \mathbb{Z} \cong L \neq 0$.

Corolario 4.6 *Si M y L son k-módulos libres con bases $\{u_1, ..., u_m\}$ y $\{v_1, ..., v_n\}$ respectivamente, entonces $M \otimes L$ también es libre, de rango mn; y admite como base al conjunto $\{u_i \otimes v_j \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$.*

Demostración: Tenemos que $M = \bigoplus_{i=1}^{m} ku_i$ y $L = \bigoplus_{j=1}^{n} kv_j$, por lo que $M \otimes L \cong \bigoplus_{j=1}^{n} M \otimes (kv_j) \cong \bigoplus_{i,j=1}^{m,n} k(u_i \otimes v_j)$, usando las afirmaciones 1,3 y 4 de la proposición. □

Corolario 4.7 *Sean k un anillo conmutativo, $\mathfrak{a}$ un ideal y M un k-módulo, entonces $(k/\mathfrak{a}) \otimes_k M \cong M/\mathfrak{a}M$.*

Demostración: Consideremos el producto tensorial con M de la sucesión exacta $0 \to \mathfrak{a} \to k \to k/\mathfrak{a} \to 0$. Obtenemos la siguiente sucesión exacta $\mathfrak{a}\otimes M \to k\otimes M \to (k/\mathfrak{a})\otimes M \to 0$, donde $k\otimes M \cong M$ y donde la imagen de $\mathfrak{a}\otimes M$ ante este isomorfimo es $\mathfrak{a}M$. Así, $(k/\mathfrak{a})\otimes M \cong (k\otimes M)/\operatorname{Im}(\mathfrak{a}\otimes M) \cong (M/\mathfrak{a}M)$. □

Algebras tensorial, simétrico y alternante

Sea M un k-módulo. Al producto tensorial $M \otimes \cdots \otimes M$ con n factores lo escribimos $T^n_k M$, omitiendo el subíndice cuando quede entendido. Al k-módulo

$$T(M) = \coprod_{n\geq 0} T^n M \tag{4.1}$$

le damos una estructura de k-álgebra con una multiplicación k-bilineal

$$T(M) \times T(M) \to T(M) \text{ tal que } T^r M \times T^s M \to T^{r+s} M, \tag{4.2}$$

donde el producto de los generadores $a_1 \otimes \cdots \otimes a_r$ y $b_1 \otimes \cdots \otimes b_s$ se define como $a_1 \otimes \cdots \otimes a_r \otimes b_1 \otimes \cdots \otimes b_s$. Aquí, el morfismo estructural del álgebra $T(M)$ es la composición de la identidad $k \to T^0 M$ con la inclusión canónica $T^0 M \hookrightarrow \coprod_{n\geq 0} T^n M = T(M)$. El resultado es el **álgebra tensorial** $T(M)$.

Un álgebra con una descomposición como la de (4.1) que cumple con la condición (4.2) es un **álgebra graduado**.

El **álgebra simétrico** $S(M)$ del k-módulo M es el k-álgebra que resulta del cociente de $T(M)$ entre el ideal bilateral $\mathfrak{s}$ generado por las expresiones $a \otimes b - b \otimes a$, para todos $a, b \in M$. Escribimos $S^n M = T^n M/(\mathfrak{s} \cap T^n M)$ para considerar el morfismo de módulos

$$\varphi = \coprod_{n\geq 0} \varphi_n : T(M) \to \coprod_{n\geq 0} S^n M$$

obtenido de aplicar los morfismos naturales $\varphi_n : T^n M \to S^n M$ a las componentes de $T(M)$. Aquí, $\ker \varphi = \coprod_{n\geq 0}(\mathfrak{s} \cap T^n M)$. El ideal $\mathfrak{s}$ cumple con la igualdad $\mathfrak{s} = \coprod_{n\geq 0}(\mathfrak{s} \cap T^n M)$, por lo que es un **ideal homogéneo**. Esto nos permite considerar a φ como un morfismo de anillos, que le da estructura de álgebra graduado al álgebra simétrico $S(M)$:

$$S(M) = T(M)/\mathfrak{s} \cong \coprod_{n\geq 0} S^n M.$$

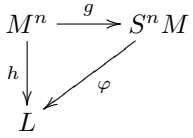
Es fácil ver que si $n \geq 2$, $\sigma \in S_n$ es una permutación y $a_1, ..., a_n \in M$, entonces $\mathfrak{s}$ contiene a $(a_1 \otimes \cdots \otimes a_n) - (a_{\sigma(1)} \otimes \cdots \otimes a_{\sigma(n)})$.

El **álgebra alternante** $\bigwedge(M)$ del k-módulo M es el k-álgebra que se obtiene al formar el cociente del álgebra tensorial $T(M)$ entre el ideal bilateral $\mathfrak{a}$ generado por las expresiones $a \otimes a$, para todo $a \in M$. Si $a, b \in M$, entonces $\mathfrak{a}$ contiene a las expresiones $a \otimes b + b \otimes a = (a+b) \otimes (a+b) - a \otimes a - b \otimes b$. Para $a_1, ..., a_r \in M$, la imagen de $a_1 \otimes \cdots \otimes a_r$ en $\bigwedge(M)$ se escribe $a_1 \wedge \cdots \wedge a_r$, de manera que $a \wedge b + b \wedge a = 0$, siempre que $a, b \in M$. Como $\mathfrak{a}$ es homogéneo, tenemos que $\bigwedge(M) = \coprod_{n \geq 0} \bigwedge^n M$, donde $\bigwedge^n M = T^n M / (\mathfrak{a} \cap T^n M)$.

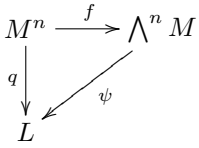
Una función $f : M^n \to k$ se llama **alternante** cuando $f(m_1, ..., m_n) = 0$ si $m_i = m_j$ para algunos $i \neq j$.

Observaciones. Sean n un entero positivo fijo y M un k-módulo. Así como $T^n M$ satisface una propiedad universal, tenemos que

1. La composición de los morfismos naturales $g : M^n \to T^n M \to S^n M$ es multilineal y simétrica; y cumple con la condición universal de que para todo k-módulo L y toda función multilineal y simétrica $h : M^n \to L$, existe un único morfismo $\varphi : S^n M \to L$ que hace conmutativo al diagrama



2. La composición de los morfismos naturales $f : M^n \to T^n M \to \bigwedge^n M$ es multilineal y alternante; y cumple con la condición universal de que para todo k-módulo L y toda función multilineal y alternante $q : M^n \to L$, existe un único morfismo $\psi : \bigwedge^n M \to L$ que hace conmutativo al diagrama



Teorema 4.8 *Sean k un anillo conmutativo, M y N k-módulos. Entonces:*

1. *Cada $T^n(M \oplus N)$ de la descomposición de $T(M \oplus N)$, es la suma directa de todos los productos tensoriales de k-módulos, con n factores, elegidos entre M y N.*

2. *$S(M \oplus N) \cong S(M) \otimes S(N)$ es un isomorfismo de álgebras, siempre que $S(M) \otimes S(N)$ esté provisto de la multiplicación conmutativa.*

3. *$\bigwedge(M \oplus N) \cong \bigwedge(M) \otimes \bigwedge(N)$ es un isomorfismo de álgebras si a $\bigwedge(M) \otimes \bigwedge(N)$ se le provee de la multiplicación bilineal que cumple $(a \otimes b) \wedge (c \otimes d) = (-1)^{pq}(ac \otimes bd)$, para todos $b \in \bigwedge^q N$, $c \in \bigwedge^p M$.*

Demostración: La primera afirmación es consecuencia de la Proposición 4.5. Por ejemplo,

$$T^2(M \oplus N) = (M \otimes M) \oplus (M \otimes N) \oplus (N \otimes M) \oplus (N \otimes N).$$

Así, podríamos escribir

$$T(M \oplus N) = T(M) \otimes T(N) \otimes T(M) \otimes T(N) \otimes \cdots$$

A partir de aquí, las dos siguientes afirmaciones admiten demostraciones análogas entre sí. Veamos por ejemplo la última.

$\bigwedge(M \oplus N)$ es el resultado de dividir $T(M \oplus N)$ entre el ideal $\mathfrak{a}$ generado por las expresiones $a \otimes a$, para todos $a \in M \oplus N$. Sea $\mathfrak{b}$ el subideal de $\mathfrak{a}$ generado por las expresiones $a \otimes b + b \otimes a$ para todos $a \in M,\ b \in N$; y sea $\overline{\mathfrak{a}}$ la imagen de $\mathfrak{a}$ en $T(M \oplus N)/\mathfrak{b}$. Entonces, $T(M \oplus N)/\mathfrak{b} \cong T(M) \otimes T(N)$, mientras que

$$\bigwedge(M \oplus N) \cong [T(M) \otimes T(N)]/\overline{\mathfrak{a}} \cong \bigwedge(M) \otimes \bigwedge(N).$$

Este último objeto provisto de la multiplicación indicada. □

Corolario 4.9 *Sea M un k-módulo libre con base $\{u_1, ..., u_m\}$. Entonces:*

1. *$T(M)$ y todos los $T^n M$ también son libres. El rango de $T^n M$ es m^n; y $\{u_{i_1} \otimes \cdots \otimes u_{i_n} \mid 1 \leq i_1, ..., i_n \leq m\}$ es una base de $T^n M$, por lo que $T(M)$ se identifica con el álgebra no comutativo de polinomios en las variables $\{u_1, ..., u_m\}$ con coeficientes en k.*

2. *$S(M)$ y todos los $S^n M$ también son libres. El rango de $S^n M$ es $\binom{m+n-1}{m-1}$; y el conjunto de monomios de grado n en $u_1, ..., u_m$ es una base de $S^n M$, por lo que $S(M)$ se identifica con el álgebra de polinomios en esas variables con coeficientes en k.*

3. *$\bigwedge(M)$ y todos los $\bigwedge^n M$ también son libres. El rango de $\bigwedge^n M$ es $\binom{m}{n}$; y $\{u_{i_1} \wedge \cdots \wedge u_{i_n} \mid 1 \leq i_1 < \cdots < i_n \leq m\}$ es una base de $\bigwedge^n M$.*

Demostración: Por el Corolario 4.6, tenemos que $T(M)$ y todo $T^n M$ también son libres, como en el enunciado. Así, $T(M)$ se identifica con el álgebra no comutativo de polinomios en las u_i con coeficientes en k.

Después, inducción en m produce $M = k^m = ku_1 \oplus \coprod_{i=2}^m ku_i$. Aquí,

$$S(M) = S(ku_1) \otimes S(\coprod_{i=2}^m ku_i),$$

donde $S(ku_1)$ tiene como base a las potencias de u_1, mientras que la hipótesis inductiva dice que $S(\coprod_{i=2}^m ku_i)$ es el anillo de polinomios en $u_2, ..., u_m$. La conclusión se obtiene de ahí usando el Teorema 4.8.2.

Finalmente, $M = ku_1 \oplus \coprod_{i=2}^{m} ku_i$ produce $\bigwedge M = \bigwedge ku_1 \otimes \bigwedge \coprod_{i=2}^{m} ku_i$ con $\bigwedge(ku_1) = k[u_1]/(u_1^2) = k \oplus ku_1$. De manera que

$$\bigwedge(M) = (k \oplus ku_1) \otimes \bigwedge(\coprod_{i=2}^{m} ku_i),$$
$$\bigwedge^{n}(k^m) = \bigwedge^{n}(k^{m-1}) \oplus [ku_1 \otimes \bigwedge^{n-1}(k^{m-1})],$$

de donde se obtiene la conclusión usando el Teorema 4.8.3. □

Transformaciones de rango uno

Supongamos que M y L son espacios vectoriales sobre un campo k, con bases $\{u_1, ..., u_m\}$ y $\{v_1, ..., v_n\}$ respectivamente. Sabemos que entonces el dual $M^* = \mathrm{Hom}_k(M, k)$ también es libre, con base $\{u_1^*, ..., u_m^*\}$, donde $u_i^*(u_j) = \delta_{ij}$. El Corolario 4.6, garantiza que $L \otimes M^*$ es libre, con base

$$\{v_j \otimes u_i^* \mid 1 \leq j \leq n,\ 1 \leq i \leq m\}.$$

Así, $L \otimes M^* \cong \mathrm{Hom}_k(M, L)$, al ser ambos módulos libres de rango mn; pero este isomorfismo es "natural", por lo que no le asignamos ningún símbolo, simplemente consideramos que $v \otimes u^*$ pertenece a $\mathrm{Hom}_k(M, L)$; y que actúa así: $(v \otimes u^*)(x) = u^*(x)v$, para todos $x \in M,\ u^* \in M^*$ y $v \in L$.

Observando que $(v_j \otimes u_i^*)(u_p) = \delta_{ip} v_j$, es claro que a T le corresponde la matriz $A = (a_{ij}) \in M_{m \times n}$ con respecto a las bases elegidas si y sólo si $T = \sum_{ij} a_{ij}(v_j \otimes u_i^*)$.

El **rango** de una transformación $T : M \to L$ es el de su imagen $T(M)$; así, cada $v \otimes u^* \in \mathrm{Hom}_k(M, L)$ es de **rango uno**. Aquí tenemos que

$$\mathrm{Im}(v \otimes u^*) = \langle v \rangle \text{ y que } \ker(v \otimes u^*) = \ker u^*.$$

Observaciones. Si V es un espacio vectorial de dimensión n, entonces se tiene para todos $v, z \in V,\ u^*, w^* \in V^*$ y $T \in \mathrm{End}\, V$ que:

1. $T(v \otimes u^*) = Tv \otimes u^*$,
2. $(v \otimes u^*)T = v \otimes T^* u^*$,
3. $(v \otimes u^*)(z \otimes w^*) = (u^* z)(v \otimes w^*)$,
4. $(v \otimes u^*)^* = u^* \otimes v^{**} \cong u^* \otimes v$.

Teorema 4.10 *a) Si* $\mathrm{rango}\, T = r$, *entonces* T *es la suma de* r *transformaciones de rango uno; pero no es suma de ningún numero menor de tales transformaciones.*

b) $\mathrm{rango}\, T = \mathrm{rango}\, T^*$.

Demostración: *a*) Sean $\{w_1, ..., w_r\}$ una base de $T(M)$ y $v \in M$, entonces existen escalares $c_i \in k$ para $1 \leq i \leq r$ tales que $T(v) = \sum_{i=1}^{r} c_i w_i$; y que definen funciones lineales $f_i \in M^*$ con $f_i(v) = c_i$. Así resulta que $T = \sum_{i=1}^{r}(w_i \otimes f_i)$.

Supongamos que $T = \sum_{i=1}^{s} B_i$, con rango $B_i = 1$, para toda B_i, entonces $T(M) \subseteq \sum_{i=1}^{s} B_i(M)$, por lo que $r = \text{rango}\, T(M) \leq \sum_{i=1}^{s} \text{rango}\, B_i = s$.

b) Supongamos que rango $T = r$, de manera que existe $\{v_1, ..., v_r\}$ linealmente independiente con

$$T = \sum_{i=1}^{r}(v_i \otimes u_i^*) \Rightarrow T^* = \sum_{i=1}^{r}(v_i \otimes u_i^*)^* = \sum_{i=1}^{r}(u_i^* \otimes v_i),$$

identificando a cada v_i^{**} con v_i, por lo que rango $T^* \leq r$.

Afirmamos que $\{u_1^*, ..., u_r^*\}$ es linealmente independiente. De no ser así, tal vez reordenando los índices, se tendrían expresiones $u_i^* = \sum_{j=1}^{s} a_{ij} u_j^*$ para $1 \leq i \leq r$, con escalares a_{ij}, donde $s < r$. Pero entonces tendríamos

$$T = \sum_{i=1}^{r}(v_i \otimes \sum_{j=1}^{s} a_{ij} u_j^*) = \sum_{j=1}^{s}(\sum_{i=1}^{r} a_{ij} v_i) \otimes u_j^*),$$

una suma de s transformaciones de rango ≤ 1; y una contradicción.

Elegimos $w_j^* \in L^*$ tales que $w_j^*(v_i) = \delta_{ij}$, para tener $T^*(w_j^*) = u_j^*$. Así, $u_i^* \in \text{Im}\, T^*$, para todo $1 \leq i \leq r$; y rango $T^* = r$. □

Dados un campo k y una matriz $A \in M_{m \times n}(k)$, definimos el **rango renglón** de A como la dimensión del subespacio de $k^{(m)}$ generado por los renglones de A. El **rango columna** de A es la dimensión del subespacio de $k^{(n)}$ generado por las columnas de A.

Corolario 4.11 rango renglón A = rango columna A.

Demostración: Supongamos que la matriz A está asociada a la función $T : k^{(m)} \to k^{(n)}$, entonces A^t está asociada a $T^* : k^{(n)} \to k^{(m)}$. La conclusión es consecuencia de que rango renglón A = rango T y que rango columna A = rango T^*. □

Teorema 4.12 *Si $T : V \to W$ es una transformación lineal y V es de dimensión finita, entonces* $\dim V = \text{rango}\, T + \dim \ker T$.

Demostración: Sea $\{u_1, ..., u_r\}$ una base de $\ker T$ que se extiende a la base $\{u_1, ..., u_r, v_1, ..., v_s\}$ de V. Aquí tenemos que $\dim \ker T = r$, que $\dim V = r + s$ y que $B = \{T(v_1), ..., T(v_s)\}$ genera a $T(V)$; pero B resulta ser una base de $T(V)$ y $s = \circ(B) = \text{rango}\, T$, al ser B linealmente independiente:

$$\sum_{i=1}^{s} c_i T(v_i) = 0 \Rightarrow T(\sum_{i=1}^{s} c_i v_i) = 0 \Rightarrow \sum_{i=1}^{s} c_i v_i \in \ker T \Rightarrow c_i = 0, \ \forall\ i. \ \square$$

Ejercicios

1. *a*) Sea G un grupo finito. Demuestre que $\circ(G)$ es la dimensión como espacio vectorial sobre un campo k del k-álgebra de grupo $k[G]$.

   b) Sea H el grupo de cuaternios de la Sección 1.3. Investigue si son isomorfos $\mathbb{R}[\mathsf{H}]$ y el álgebra de cuaternios reales $\mathbb{H}$.

2. Sea $f : k \to A$ un k-álgebra, libre como k-módulo, con $\operatorname{rango}_k A = n$. Demuestre que A es isomorfo con un subálgebra de $\operatorname{End}_k k^{(n)}$.

3. **La multiplicación de Arens.** Sea V un espacio vectorial sobre k de dimensión infinita, provisto de una multiplicación bilineal $V \times V \to V$. Definimos una serie de multiplicaciones como se indica.

$$\begin{aligned} V^* \times V \to V^*, \quad & (u^* \cdot v)(w) = u^*(vw), \\ V^{**} \times V^* \to V^*, \quad & (u^{**} \cdot v^*)(w) = u^{**}(v^* \cdot w), \\ V^{**} \times V^{**} \to V^{**}, \quad & (u^{**} \cdot v^{**})(w^*) = u^{**}(v^{**} \cdot w^*). \end{aligned}$$

   Demuestre que $h : V \to V^{**}$ dada por $h(u)(v^*) = v^*(u) \in k$, para $u \in V, v^* \in V^*$ es inyectiva. Demuestre que $V^{**} \times V^{**} \to V^{**}$ extiende a la multiplicación original y que es asociativa, si la multiplicación original lo era.

4. Verifique que la **suma directa** $\coprod_{i \in I} M_i$ de k-módulos satisface la siguiente propiedad universal: Dados un k-módulo C y k-morfismos $\{g_i : M_i \to C \mid i \in I\}$; existe un único k-morfismo $g : \coprod_{i \in I} M_i \to C$ tal que $g \circ j_i = g_i$, para todo $i \in I$.

5. Sean R un anillo local, $\mathfrak{m}$ su ideal máximo, $k = R/\mathfrak{m}$ y M un R-módulo finitamente generado. Demuestre que $k \otimes M = 0 \Rightarrow M = 0$.

6. Sean k un campo y V un espacio vectorial sobre k de dimensión finita. Demuestre que para $T, S \in \operatorname{End} V$, es cierto que

   (a) $\operatorname{rango}(T + S) \leq \operatorname{rango}(T) + \operatorname{rango}(S)$.

   (b) $\operatorname{rango}(T \circ S) \leq \min\{\operatorname{rango}(T), \operatorname{rango}(S)\}$.

   (c) S invertible $\Rightarrow \operatorname{rango}(T \circ S) = \operatorname{rango}(T) = \operatorname{rango}(S \circ T)$.

7. Sean k un campo, V un espacio vectorial sobre k de dimensión finita y W un subespacio de V. Demuestre que $\operatorname{End} V$ es simple y que

   (a) $\{T \in \operatorname{End} V \mid \operatorname{Im} T \subseteq W\}$ es un ideal derecho de $\operatorname{End} V$.

   (b) Todo ideal derecho de $\operatorname{End} V$ es de la forma anterior. (Sugerencia: Reduzca al caso de transformaciones de rango uno).

   (c) $\{T \in \operatorname{End} V \mid W \subseteq \ker T\}$ es un ideal izquierdo de $\operatorname{End} V$.

   (d) Todo ideal izquierdo de $\operatorname{End} V$ es de la forma anterior.

   (e) Determine los ideales izquierdos y derechos mínimos $\neq 0$ de $\operatorname{End} V$.

4.3 Determinantes

Sean k un anillo conmutativo y V un k-módulo libre con base $\{\epsilon_1, ..., \epsilon_n\}$. En la descomposición $\bigwedge V = \bigwedge^0 V \oplus \bigwedge^1 V \oplus \bigwedge^2 V \oplus \cdots \oplus \bigwedge^n V$, se tiene que cada $\bigwedge^r V$ es un módulo libre con base $\{\epsilon_{i_1} \wedge \cdots \wedge \epsilon_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq n\}$. En particular, $\bigwedge^0 V \cong k$, $\bigwedge^1 V \cong V$ y $\dim \bigwedge^n V = 1$. Además, $\bigwedge^n V$ tiene como base al conjunto con un sólo elemento $\{\epsilon_1 \wedge \cdots \wedge \epsilon_n\}$.

Dados n vectores $u_1, ..., u_n \in V$ con $u_i = \sum_{j=1}^n a_{ij}\epsilon_j$ para cada $1 \leq i \leq n$, calculamos su producto exterior, para obtener

$$u_1 \wedge \cdots \wedge u_n = \sum_{j=1}^n a_{1j}\epsilon_j \wedge \cdots \wedge \sum_{j=1}^n a_{nj}\epsilon_j = \Delta\ \epsilon_1 \wedge \cdots \wedge \epsilon_n,$$

para algún elemento $\Delta \in k$. Esto nos permite definir una función multilineal y alternante $\det : V^n \to k$ llamada **determinante**. De manera que $\det(u_1, ..., u_n) = \Delta$.

Dada $A = (a_{ij}) \in M_n(k)$, una matriz $n \times n$; definimos el **determinante de** A, escrito $\det A$, como $\det(u_1, ..., u_n)$, donde $u_i = \sum_{j=1}^n a_{ij}\epsilon_j$ son los vectores renglón de A. Así, tenemos otra función, también llamada determinante, $\det : M_n(k) \to k$. También escribimos $\det A$ así:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Observaciones. Obtenemos las siguientes propiedades de los determinantes a partir de su definición.

1. El caso 2×2.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \quad \text{pues}$$
$$(a_{11}\epsilon_1 + a_{12}\epsilon_2) \wedge (a_{21}\epsilon_1 + a_{22}\epsilon_2) = (a_{11}a_{22} - a_{12}a_{21})\ (\epsilon_1 \wedge \epsilon_2).$$

2. El determinante de una matriz diagonal es

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \cdots a_{nn},$$

pues $(a_{11}\epsilon_1) \wedge \cdots \wedge (a_{nn}\epsilon_n) = (a_{11} \cdots a_{nn})\ \epsilon_1 \wedge \cdots \wedge \epsilon_n$. Casos particulares importantes son el determinante de la matriz identidad, $\det(\delta_{ij}) = 1$; y el determinante de una matriz escalar, $\det(b\delta_{ij}) = b^n$.

3. Efectuando la multiplicación exterior para una matriz $A = (a_{ij})$, vemos que
$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} a_{i,\sigma(i)}. \tag{4.3}$$
De manera que $\det : M_n(k) \to k$ es una función polinomial homogénea de grado n con $n!$ términos, la mitad de los cuales tienen coeficiente 1; y la otra mitad tienen coeficiente -1. Cada término de (4.3) incluye exactamente un factor de cada renglón (columna) de A.

4. El determinante de una matriz es igual al de su transpuesta.
Sean $A = (a_{ij})$ y $B = A^t = (b_{ij})$, de manera que $b_{ij} = a_{ji}$, para cualesquiera índices i, j. Escribimos $\tau = \sigma^{-1}$, para obtener
$$\begin{aligned} \det A^t = \det B &= \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} b_{i,\sigma(i)} = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} a_{\sigma(i),i} \\ &= \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} a_{i,\sigma^{-1}(i)} = \sum_{\tau \in S_n} (-1)^\tau \prod_{i=1}^{n} a_{i,\tau(i)} = \det A. \end{aligned}$$

5. Dada $A \in M_n(k)$, con vectores columna $v_p = \sum_{i=1}^{n} a_{ip}\epsilon_i$, también podemos considerar $\det(v_1, ..., v_n)$. Una consecuencia inmediata de la observación anterior es que $\det A = \det(v_1, ..., v_n)$. De manera que podemos decir que el determinante de una matriz es una función multilineal y alternante de sus renglones o de sus columnas.

Teorema 4.13 *Si $A, B \in M_n(k)$, entonces* $\det(AB) = (\det A)(\det B)$.

Demostración: Sean $u_1, ..., u_n$ los vectores renglón de B y $A = (a_{ij})$, entonces los renglones de AB son $\sum_{j=1}^{n} a_{1j}u_j, ..., \sum_{j=1}^{n} a_{nj}u_j$. El cálculo
$$\begin{aligned} \sum_{j=1}^{n} a_{1j}u_j \wedge \cdots \wedge \sum_{j=1}^{n} a_{nj}u_j &= (\det A)\ u_1 \wedge \cdots \wedge u_n \\ &= (\det A)(\det B)\ \epsilon_1 \wedge \cdots \wedge \epsilon_n \end{aligned}$$
demuestra que $\det(AB) = (\det A)(\det B)$. □

Teorema 4.14 (Regla de Cramer) *Si el sistema de ecuaciones lineales*
$$\begin{array}{ccccccc} a_{11}\ X_1 & + & \cdots & + & a_{1n}\ X_n & = & c_1 \\ & & \ddots & & & & \vdots \\ a_{n1}\ X_1 & + & \cdots & + & a_{nn}\ X_n & = & c_n \end{array} \tag{4.4}$$
satisface la condición $D = \det(a_{ij}) \in k^\star$, entonces existen soluciones únicas dadas por: $X_i = D^{-1} \det(a_1, ..., a_{i-1}, c, a_{i+1}, ..., a_n)$, para $1 \leq i \leq n$, donde c es el vector con coordendas $c_1, ..., c_n$; mientras que a_j es el j-ésimo vector columna de la matriz $A = (a_{ij})$.

Demostración: El sistema de ecuaciones (4.4) puede expresarse como la igualdad de matrices en $M_{n\times 1}(k[X_1,...,X_n])$:

$$(\,a_1 \quad \cdots \quad a_n\,)\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Así, (4.4) dice que $\sum_{j=1}^n X_j a_j = c$. La conclusión es consecuencia del cálculo siguiente

$$\begin{aligned} &\det(a_1,...,a_{i-1},c,a_{i+1},...,a_n) \\ &= \det(a_1,...,a_{i-1},\sum_{j=1}^n X_j a_j, a_{i+1},...,a_n) \\ &= \det(a_1,...,a_{i-1},X_i a_i,a_{i+1},...,a_n) = X_i D. \ \square \end{aligned}$$

Teorema 4.15 *Dado un anillo conmutativo k, una matriz $A \in M_n(k)$ es invertible si y sólo si su determinante lo es; en cuyo caso,* $\det(A^{-1}) = (\det A)^{-1}$.

Demostración: Veamos primero que A invertible $\Rightarrow \det A \in k^\star$:

$$AA^{-1} = 1 \Rightarrow (\det A)(\det A^{-1}) = 1 \Rightarrow \det(A^{-1}) = (\det A)^{-1}.$$

Recíprocamente, si $\det A \in k^\star$, entonces cada sistema de ecuaciones lineales $\sum_{j=1}^n X_j a_j = \epsilon_i^t$ con $1 \le i \le n$, formado usando las columnas a_j de A, tiene como solución única al vector columna $(b_{1i},...,b_{ni})^t$. Entonces la matriz $B = (b_{ij})$ satisface $AB = 1$. Así, $B = A^{-1}$. $\square$

Observación. Si k es un campo, entonces el Teorema 4.15 afirma que A es invertible si y sólo si $\det A \neq 0$. La Proposición 4.2 implica que los renglones de A forman un conjunto linealmente independiente si y sólo si $\det A \neq 0$.

Dados un anillo conmutativo k y una matriz $A = (a_{ij}) \in M_n(k)$, escribimos $N = \{1,...,n\}$; y consideramos los renglones de A:

$$a_i = \sum_{j=1}^n a_{ij}\epsilon_j, \text{ para } 1 \le i \le n.$$

La definición de $\det A = \Delta(A)$ fue a través de la igualdad

$$a_1 \wedge \cdots \wedge a_n = \Delta(A)\ \epsilon_1 \wedge \cdots \wedge \epsilon_n.$$

Elijamos ahora un subconjunto $P = \{i_1,...,i_p\} \subseteq N$, donde $i_1 < \cdots < i_p$. Obtenemos una expresión única

$$a_{i_1} \wedge \cdots \wedge a_{i_p} = \sum_{Q\subseteq N, \circ(Q)=p} \Delta_P^Q(A)\ \epsilon_{j_1} \wedge \cdots \wedge \epsilon_{j_p}, \tag{4.5}$$

donde $\Delta_P^Q(A) \in k$, para cada $Q = \{j_1, ..., j_p\} \subseteq N$ con $\circ(Q) = p$. Aquí suponemos que $j_1 < \cdots < j_p$.

Los elementos $\Delta_P^Q(A)$, también escritos $\Delta_{i_1,...,i_p}^{j_1,...,j_p}(A)$, son los **menores** $p \times p$ de A. El número de tales menores con P fijo es $\binom{n}{p}$.

Para calcular cada $\Delta_P^Q(A)$, consideramos las proyecciones $f_Q : V \to V_Q$ en los distintos submódulos $V_Q = \langle \epsilon_j \mid j \in Q \rangle$, de manera que

$$\begin{aligned} f_Q(a_i) &= a_{ij_1}\epsilon_{j_1} + \cdots + a_{ij_p}\epsilon_{j_p}, \\ f_Q(a_{i_1}) \wedge \cdots \wedge f_Q(a_{i_p}) &= \Delta_P^Q(A)\ \epsilon_{j_1} \wedge \cdots \wedge \epsilon_{j_p}. \end{aligned} \tag{4.6}$$

La ecuación (4.6) nos aclara que $\Delta_P^Q(A)$ es el determinante de la submatriz de A formada por los renglones en P y las columnas en Q. Observamos que $\Delta_N^N(A) = \Delta(A)$, convenimos que $\Delta_\varnothing^\varnothing(A) = 1$ y escribimos $P' = N \smallsetminus P$.

Teorema 4.16 (Expansión de Laplace) *Dada $A = (a_{ij}) \in M_n(k)$, elegimos un conjunto de renglones $P = \{i_1, ..., i_p\}$ con $i_1 < \cdots < i_p$.*

Entonces

$$\det A = \sum_{Q \subseteq N, \circ(Q) = p} (-1)^{(i_1 + j_1) + \cdots + (i_p + j_p)} \Delta_P^Q(A) \Delta_{P'}^{Q'}(A), \tag{4.7}$$

donde $Q = \{j_1, ..., j_p\}$ con $j_1 < \cdots < j_p$.

Demostración: Efectuamos el cálculo

$$\begin{aligned} (-1)^{(i_1 - 1) + \cdots + (i_p - p)}(a_1 \wedge \cdots \wedge a_n) &= \\ (a_{i_1} \wedge \cdots \wedge a_{i_p}) \wedge (a_{i_{p+1}} \wedge \cdots \wedge a_{i_n}) &= \\ \sum_{Q \subseteq N, \circ(Q) = p} \Delta_P^Q(A)\ \epsilon_{j_1} \wedge \cdots \wedge \epsilon_{j_p} \wedge (a_{i_{p+1}} \wedge \cdots \wedge a_{i_n}) &= \\ \sum_{Q \subseteq N, \circ(Q) = p} \Delta_P^Q(A)\ \epsilon_{j_1} \wedge \cdots \wedge \epsilon_{j_p} \wedge (\Delta_{P'}^{Q'}(A)\ \epsilon_{j_{p+1}} \wedge \cdots \wedge \epsilon_{j_n}) &= \\ \sum_{Q \subseteq N, \circ(Q) = p} (-1)^{(j_1 - 1) + \cdots + (j_p - p)} \Delta_P^Q(A)\ \Delta_{P'}^{Q'}(A)\ \epsilon_1 \wedge \cdots \wedge \epsilon_n, & \end{aligned}$$

donde $P' = \{i_{p+1}, ..., i_n\}$ con $i_{p+1} < \cdots < i_n$ y $Q' = \{j_{p+1}, ..., j_n\}$ con $j_{p+1} < \cdots < j_n$, de manera que $a_1 \wedge \cdots \wedge a_n =$

$$\sum_{Q \subseteq N, \circ(Q) = p} (-1)^{(i_1 + j_1) + \cdots + (i_p + j_p)} \Delta_P^Q(A)\ \Delta_{P'}^{Q'}(A)\ \epsilon_1 \wedge \cdots \wedge \epsilon_n,$$

de donde se obtiene la conclusión. □

Es posible enunciar un resultado análogo partiendo de un conjunto fijo de columnas. En vista de (4.7), se dice que

$$(-1)^{(\sum_{i \in P} i + \sum_{j \in Q} j)}\ \ \Delta_{P'}^{Q'}(A)$$

es el **cofactor** de $\Delta_P^Q(A)$.

En el caso en que $P = \{i\}$ y $Q = \{j\}$, se tiene que $\Delta_P^Q(A) = a_{ij}$ y que su cofactor, escrito A_{ij} es $(-1)^{i+j}$ por el determinante de la submatriz de A obtenida eliminando el renglón i y la columna j.

La expresión

$$\det A = a_{i1}A_{i1} + \cdots + a_{in}A_{in}$$

es la **expansión a lo largo del renglón** i.

La matriz $\operatorname{cof} A = (A_{ij})$ es la **matriz de cofactores** de A. La matriz $\operatorname{adj} A = (A_{ij})^t$ es la **matriz adjunta** de A.

Teorema 4.17 *Para $A = (a_{ij}) \in M_n(k)$, se tiene que*

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = (\det A)1,$$

donde 1 *es la matriz identidad $n \times n$.*

Demostración: El elemento en la posición ij de $A(\operatorname{adj} A)$ es

$$a_{i1}A_{j1} + \cdots + a_{in}A_{jn};$$

pero tenemos que $a_{i1}A_{i1} + \cdots + a_{in}A_{in} = \det A$, mientras que si $i \neq j$, se tiene que $a_{i1}A_{j1} + \cdots + a_{in}A_{jn} = 0$, al ser la expansión a lo largo del renglón j de la matriz con dos renglones iguales, obtenida de A reemplazando al renglón j por el renglón i. □

Corolario 4.18 *Si $A \in M_n(k)$ y $\det A \in k^\star$, entonces*

$$A^{-1} = (\det A)^{-1}(\operatorname{adj} A).$$

Ejemplo. En el caso 2×2, se tiene que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = ad - bc = \Delta,$$

de manera que si existe Δ^{-1}, entonces

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = (ad - bc)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Supongamos que k es un campo y que $A = (a_{ij}) \in M_{m\times n}(k)$. Decimos que A tiene **rango determinantal** r en caso de que exista un menor $r \times r$ de A distinto de cero; pero con todos los menores $(r+1) \times (r+1)$ de A iguales a cero.

Teorema 4.19 *Si k es un campo y $A = (a_{ij}) \in M_{m\times n}(k)$, entonces* $\operatorname{rango} A = \operatorname{rango}$ determinantal A.

Demostración: Procedemos por inducción en $r = \text{rango } A$. Aquí existe un conjunto linealmente independiente de r renglones de A; y todo conjunto que contenga más de r renglones de A es linealmente dependiente.

Todo menor de A de tamaño $s \times s$ con $s > r$ vale cero, pues la submatriz asociada consiste de las coordenadas de una proyección de s vectores renglón de A, que forman un conjunto linealmente dependiente.

Supongamos que los primeros r renglones de A son linealmente independientes. Sea B la submatriz de A formada por estos renglones.

Como el conjunto de los primeros $r-1$ renglones de A (ó de B) es también linealmente independiente, la hipótesis inductiva garantiza que existe un menor de B, que excluya a su último renglón, de tamaño $(r-1) \times (r-1)$ distinto de cero. Supongamos que $\Delta^{1,\ldots,r-1}_{1,\ldots,r-1}(A) = \Delta^{1,\ldots,r-1}_{1,\ldots,r-1}(B) \neq 0$.

Queremos demostrar que existe un menor $r \times r$ de B distinto de cero. Supongamos que esto es falso. La expansión en la última columna de $\Delta^{1,\ldots,r}_{1,\ldots,r}(B)$ es:

$$\sum_{i=1}^{r} a_{ir} A_{ir} = 0, \tag{4.8}$$

donde A_{ir} es el cofactor de a_{ir} en $\Delta^{1,\ldots,r}_{1,\ldots,r}(B)$. Para cada $p < r$,

$$\sum_{i=1}^{r} a_{ip} A_{ir} = 0, \tag{4.9}$$

por ser la expansión de un determinante con dos columnas iguales. Esta igualdad también es válida para $p > r$, al ser la expansión de $\Delta^{1,\ldots,r-1,p}_{1,\ldots,r}(B)$ en la columna p.

Si $a_1, \ldots, a_r$ son los renglones de B, las ecuaciones (4.8-9) afirman que

$$\sum_{i=1}^{r} A_{ir} a_i = 0.$$

Esta es una contradictoria relación de dependencia lineal, pues $A_{rr} = \Delta^{1,\ldots,r-1}_{1,\ldots,r-1}(B) \neq 0$. □

Teorema 4.20 *Sean k un dominio de factorización única, $R = k[X_{ij}]$ el anillo de polinomios en las n^2 variables X_{ij}, para $1 \leq i, j \leq n$; y sea $A = (X_{ij}) \in M_n(R)$. Entonces* $\det A$ *es irreducible en R.*

Demostración: El Teorema 2.54 garantiza que R también es de factorización única. Procedemos por inducción en n. Supongamos que $\det A$ admite una factorización no trivial.

Sea Q el anillo de polinomios sobre k en las variables $X_{ij} \neq X_{11}$, de manera que $R = Q[X_{11}]$. La expansión a lo largo del renglón 1 de $\det A$ es

$$\det A = \sum_{i=1}^{n} X_{1i} A_{1i} = X_{11} A_{11} + p, \text{ con } p \in Q.$$

Aquí podemos suponer que A_{11} es irreducible en Q, por lo que A_{11} también es irreducible en R; y $\det A$ sólo puede factorizarse así en $Q[X_{11}]$:

$$\det A = (aX_{11} + b)c, \text{ con } a, b, c \in Q.$$

Esto implica que $ac = A_{11}$, por lo que podemos suponer que $a = 1$ ó bien que $c = 1$. Veamos qué sucede en cada caso. Si $c = 1$, entonces la factorización es trivial; y terminamos.

La factorización no trivial de $\det A$ implica que $a = 1$; pero entonces $c = A_{11}$; y así $A_{11} \mid (\det A)$. Análogamente, $(A_{11} \cdots A_{nn}) \mid (\det A)$, que es absurdo para $n = 2$ porque $(X_{11}X_{22}) \nmid (X_{11}X_{22} - X_{12}X_{21})$; y que también es absurdo para $n > 2$ porque el grado de $(A_{11} \cdots A_{nn})$ es $n(n-1)$, mayor que el grado n de $\det A$. □

Teorema 4.21 *Sean k un campo infinito, $R = k[X_{ij}]$ el anillo de polinomios en las n^2 variables X_{ij}, para $1 \leq i, j \leq n$; y sea $f(X_{ij}) \in R$ homogéneo tal que al evaluarlo en elementos de $M_n(k)$ se tengan $f(1) = 1$ y $f(AB) = f(A)f(B)$. Entonces $f(X_{ij})$ es una potencia del polinomio* $\det(X_{ij})$.

Demostración: Supongamos que el grado de $f(X_{ij})$ es r. Como

$$(\operatorname{adj} A)A = (\det A)1 \Rightarrow f(\operatorname{adj} A)f(A) = f[(\det A)1] = (\det A)^r;$$

y dado que k es infinito, tenemos que $f(\operatorname{adj} X)f(X) = (\det X)^r$, por el Ejercicio 2.7.9, al escribir $X = (X_{ij}) \in M_n(R)$. Así, $f(X_{ij}) \mid \det(X_{ij})^r$. El teorema anterior garantiza que $f(X_{ij})$ es una potencia de $\det(X_{ij})$, porque $f(1) = 1$. □

Ejemplo. Sea k un campo tal que $p \nmid n$, donde $p = \operatorname{caract} k$. Dados $a_0, a_1, ..., a_{n-1} \in k$, el **circulante** $C(a_0, a_1, ..., a_{n-1})$ es el siguiente determinante:

$$\begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{vmatrix}$$

Supongamos que k contiene una raíz n-ésima primitiva de la unidad $\zeta = \zeta_1$ y que $\zeta_i = \zeta^i$ para $1 \leq i \leq n$. Consideremos un k-álgebra $\mathcal{A}$ con base $\{1, u, u^2, ..., u^{n-1}\}$ y con la multiplicación indicada, sujeta a la condición $u^n = 1$.

Sea $a = a_0+a_1u+a_2u^2+\cdots+a_{n-1}u^{n-1} \in \mathcal{A}$. La multiplicación izquierda por a tiene asociada la matriz

$$B = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix},$$

que satisface $\det B = C(a_0, a_1, ..., a_{n-1})$.

Sean v_i los vectores dados por

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \zeta_1 & \cdots & \zeta_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_{n-1} & \cdots & \zeta_{n-1}^{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ u \\ \vdots \\ u^{n-1} \end{pmatrix}.$$

El conjunto $\{v_1, ..., v_n\}$ es una base de $\mathcal{A}$ porque la matriz de coeficientes tiene determinante $\pm\prod_{i<j}(\zeta_i < \zeta_j) \neq 0$. La matriz asociada a la multiplicación izquierda por u con respecto a la nueva base es diagonal, dado que $uv_i = u(1 + \zeta_{i-1}u + \cdots + \zeta_{i-1}^{n-1}u^{n-1}) = \zeta^{1-i}v_i$, para todo $1 \leq i \leq n$. Por tanto, multiplicación izquierda por a tiene matriz diagonal con $a_0 + a_1\zeta^{1-i} + a_2\zeta^{2(1-i)} + \cdots + a_{n-1}\zeta^{(n-1)(1-i)}$ en la posición ii. Concluimos que $C(a_0, a_1, ..., a_{n-1}) = \prod_{i=1}^{n}(a_0 + a_1\zeta_i + a_2\zeta_i^2 + \cdots + a_{n-1}\zeta_i^{n-1})$.

Números duales y derivación de determinantes

Dado un anillo conmutativo k, definimos el álgebra de números duales sobre k como $k[\epsilon] = k[T]/(T^2)$, de manera que los elementos de $k[\epsilon]$ son de la forma $a + b\epsilon$, con $a, b \in k$. La relación $\epsilon^2 = 0$ tiene el propósito de crear una extensión natural en $k[\epsilon][X]$ para cada $f(X) \in k[X]$:

$$f(X + \epsilon) = f(X) + D[f(X)]\epsilon,$$

donde $D[f(X)]$ es la derivada de f, igualdad que puede verificarse fácilmente para monomios y extenderse linealmente al caso general.

Si consideramos una matriz $B = (b_{ij}(X)) \in M_n(k[X])$, con renglones b_i, para $1 \leq i \leq n$, tendremos que

$$\begin{aligned} \det B + D(\det B)\epsilon = \det[b_{ij}(X + \epsilon)] &= \det[b_{ij}(X) + D(b_{ij})\epsilon] \\ &= \det(b_1 + D(b_1)\epsilon, ..., b_n + D(b_n)\epsilon) \\ = \det B + \sum_{i=1}^{n} \det(b_1, ..., b_{i-1}, D(b_i), b_{i+1}, ..., b_n)\epsilon, \end{aligned}$$

por lo que

$$D(\det B) = \sum_{i=1}^{n} \det(b_1, ..., b_{i-1}, D(b_i), b_{i+1}, ..., b_n).$$

Ejercicios

1. Demuestre que el determinante de una matriz con una partición y submatrices cuadradas en la diagonal es

$$\det \begin{pmatrix} A_{11} & \star & \cdots & \star \\ 0 & A_{22} & \cdots & \star \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{pmatrix} = \prod_{i=1}^{n} (\det A_{ii}).$$

2. Dado un anillo conmutativo k con 1, sea $\{\epsilon_1, ..., \epsilon_n\}$ una base de $k^{(n)}$ y sean $u_i = \sum_{j=1}^{n} a_{ij}\epsilon_j \in k^{(n)}$, para $1 \leq i \leq n$. Demuestre que todo $v \in k^{(n)}/(u_1, ..., u_n)$ satisface $(\det A)v = 0$, donde $A = (a_{ij})$.

3. Sean k un campo y $A \in M_n(k)$. Demuestre que

   (a) $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

   (b) $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2}A$.

   (c) $\operatorname{rango}(A) = n \Rightarrow \operatorname{rango}(\operatorname{adj} A) = n$.

   (d) $\operatorname{rango}(A) = (n-1) \Rightarrow \operatorname{rango}(\operatorname{adj} A) = 1$.

   (e) $\operatorname{rango}(A) < (n-1) \Rightarrow \operatorname{rango}(\operatorname{adj} A) = 0$.

4. Para los siguientes determinantes $n \times n$, demuestre que

$$\begin{vmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ & & & \ddots & \\ 1 & 1 & 1 & \cdots & 0 \end{vmatrix} = (-1)^{n-1}(n-1).$$

$$\begin{vmatrix} a+b & a & a & \cdots & a \\ a & a+b & a & \cdots & a \\ a & a & a+b & \cdots & a \\ & & & \ddots & \\ a & a & a & \cdots & a+b \end{vmatrix} = b^{n-1}(na+b).$$

5. Sea $A = (a_{ij}) \in M_n(k)$. Demuestre que

$$\det A = a_{11}A_{11} - \sum_{i,j=2}^{n} (-1)^{i+j} a_{i1}a_{1j}\Delta_{\{1,i\}'}^{\{1,j\}'}(A).$$

6. Demuestre que el álgebra de números duales es isomorfo con $\bigwedge V$, cuando $\dim V = 1$.

4.4 Matrices sobre Dominios Principales

Suponemos que k es un anillo conmutativo. En el anillo de las matrices cuadradas $M_n(k)$ definimos la matriz E_{ij} como aquella que tiene al número 1 en la posición ij y ceros en las otras posiciones. Esto lo hacemos para cualquier pareja de índices ij.

Para $i \neq j$ y $t \in k$, definimos a la **matriz elemental de primer tipo** $x_{ij}(t) = 1 + tE_{ij}$, también llamada **transvección**. El producto y el conmutador de estas matrices se comportan así:

$$x_{ij}(t)\ x_{ij}(r) = x_{ij}(t+r). \tag{4.10}$$

$$(x_{ij}(t), x_{pq}(r)) = \begin{cases} x_{iq}(tr), & \text{si } j = p,\ i \neq q \\ 1, & \text{si } j \neq p,\ i \neq q \end{cases} \tag{4.11}$$

Recordemos que el conmutador de a, b es $(a, b) = aba^{-1}b^{-1}$; y observemos que $x_{ij}(0) = 1$, mientras que $[x_{ij}(t)]^{-1} = x_{ij}(-t)$, para todos $t \in k,\ i \neq j$. De esta manera, las expresiones (4.11) calculan conmutadores donde el conjunto $\{i, j, p, q\}$ tiene al menos tres elementos.

Una **matriz elemental de segundo tipo** $D_i(t)$ es aquella con $t \in k^\star$ en la posición ii; y con las demás coordenadas iguales a las de la matriz identidad.

Para $i \neq j$, definimos a la **matriz elemental de tercer tipo** P_{ij} como aquella obtenida a partir de la matriz identidad al intercambiar sus renglones i y j.

Observaciones. Las siguientes propiedades de las matrices elementales son inmediatas:

1. Toda matriz elemental es invertible.
2. Multiplicación izquierda por $x_{ij}(t)$ tiene el efecto de agregar t veces el renglón j al renglón i.
3. Multiplicación derecha por $x_{ij}(t)$ tiene el efecto de agregar t veces la columna i a la columna j.
4. Multiplicación izquierda (derecha) por $D_i(t)$ tiene el efecto de multiplicar por t al renglón i (a la columna i).
5. Multiplicación izquierda (derecha) por P_{ij} tiene el efecto de intercambiar a los renglones (columnas) i y j.

Definimos tres tipos de **operaciones renglón elementales**: las causadas por multiplicación izquierda con matrices elementales del tipo correspondiente. De manera semejante, las **operaciones columna elementales** son causadas por multiplicación derecha con matrices elementales del tipo correspondiente.

Dados un anillo conmutativo k y un k-módulo libre N de rango r, tenemos un antiisomorfismo de anillos $\Psi : \mathrm{End}_k(N) \to M_r(k)$. El subconjunto de automorfismos de N es un grupo ante composición de funciones, que llamamos **grupo general lineal**, lo escribimos así: $\mathrm{Aut}_k(N)$, o bien $GL(N)$; y a su imagen $\Psi(\mathrm{Aut}_k(N))$ la escribimos así: $GL_r(k)$.

El **grupo especial lineal** $SL_r(k)$ es $\{A \in GL_r(k) \mid \det A = 1\}$, un subgrupo normal de $GL_r(k)$, pues es el núcleo del morfismo *det*.

A continuación enfrentamos el problema de encontrar formas canónicas ante equivalencia para matrices sobre anillos Euclideanos o principales. El siguiente teorema es un gran paso en esta dirección.

Teorema 4.22 *Sea (k, δ) un anillo Euclideano, entonces:*

a) Las matrices elementales de 1^{er} tipo generan a $SL_n(k)$.

b) Las matrices elementales de 1^o y 2^o tipos generan a $GL_n(k)$.

c) Toda matriz $A \in M_{m\times n}(k)$ puede llevarse a través de operaciones elementales de 1^o y 2^o tipos, a una de la forma

$$\mathrm{diag}(d_1, ..., d_r, 0, ...) = \begin{pmatrix} d_1 & & & & \\ & \ddots & & & \mathbf{0} \\ & & d_r & & \\ & & & 0 & \\ & \mathbf{0} & & & \ddots \end{pmatrix},$$

con $d_1 \cdots d_r \neq 0$; y con $d_i \mid d_{i+1}$ para $1 \leq i < r$.

d) Dada $A \in M_{m\times n}(k)$, existen matrices $P \in GL_m(k)$ y $Q \in GL_n(k)$ con $PAQ = \mathrm{diag}(d_1, ..., d_r, 0, ...)$.

Demostración: *a)* Dada una matriz $A \in SL_n(k)$, consideramos al conjunto $\mathcal{A}$ de las matrices en $M_n(k)$ que se pueden obtener a partir de A por medio de una serie de operaciones elementales renglón o columna de 1^{er} tipo. Observemos que todo elemento de $\mathcal{A}$ es invertible.

Sea $B = (b_{ij}) \in \mathcal{A}$ tal que $b_{i1} \neq 0$ exhiba un valor $\delta(b_{i1})$ mínimo entre todos los valores de δ evaluada en coordenadas no cero de la primera columna de matrices en $\mathcal{A}$.

En estas condiciones, $b_{i1} \mid b_{j1}$ para todo $1 \leq j \leq n$, pues de no ser así, el algoritmo euclideano produciría $0 \neq r \in k$ tal que $b_{j1} = qb_{i1} + r$ con $\delta(r) < \delta(b_{i1})$; pero r estaría en la primera columna de alguna matriz en $\mathcal{A}$.

Así, podemos suponer que $b_{j1} = 0\ \forall j \neq i$; y vemos que $b_{i1} \in k^\star$, para conseguir una nueva $B = (b_{ij}) \in \mathcal{A}$ con $b_{11} = 1$ y con $b_{j1} = 0\ \forall j \neq 1$.

Por inducción en n, transformando a la submatriz de B obtenida al eliminar el primer renglón y la primera columna, podemos exhibir una matriz en $\mathcal{A}$ triangular superior

$$\begin{pmatrix} 1 & & \star \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix},$$

para finalmente encontrar a la matriz identidad en $\mathcal{A}$.

b) Partiendo de $A \in M_n(k)$ con $\det A = c \in k^\star$, podemos escribir

$$A = \text{diag}(c, 1, ..., 1)[\text{diag}(c^{-1}, 1, ..., 1)\ A] = \text{diag}(c, 1, ..., 1)\ B,$$

con $B = \text{diag}(c^{-1}, 1, ..., 1)\ A \in SL_n(k)$.

c) El inciso *b*) implica que las matrices elementales de 3^{er} tipo se pueden expresar en términos de las de 1^o y 2^o tipos. Por tanto, aquí podemos utilizar todo tipo de operaciones elementales.

Dada $A \in M_n(k)$, consideramos al conjunto $\mathcal{A}$ de las matrices que se pueden obtener a partir de A por medio de operaciones de 1^o y 2^o tipos. Elegimos $B = (b_{ij}) \in \mathcal{A}$ con $b_{ij} \neq 0$ y con $\delta(b_{ij})$ mínimo entre los valores posibles para coordenadas de matrices en $\mathcal{A}$. Suponemos que $i = j = 1$, para poder también suponer que $b_{1j} = b_{i1} = 0,\ \forall\ i, j > 1$. Así,

$$B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

donde además $b_{11} \mid b_{ij}$ para todos $i, j > 1$. Concluimos por inducción, pues

$$\begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

llega a la forma indicada con coordenadas que son múltiplos de b_{11}.

d) Esto es claro. □

Observaciones. Para saber que las matrices de la forma dada en *c*) son canónicas, falta establecer la unicidad de los elementos d_i. Esto lo hacemos en el siguiente teorema, para el caso más general de dominios principales, donde primero vemos la existencia de esas matrices "diagonales". Las afirmaciones *a*) y *b*) anteriores no son válidas para todo dominio principal.

Teorema 4.23 *a*) *Toda matriz $A \in M_{m\times n}(k)$ con k principal es equivalente a una de la forma*

$$\text{diag}(d_1, ..., d_r, 0, ...) = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \mathbf{0} \\ & & d_r & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \end{pmatrix},$$

con $d_1 \cdots d_r \neq 0$; y con $d_i \mid d_{i+1}$ para $1 \leq i < r$.

b) *Cada d_i es único módulo asociados; y si Δ_i es el* m.c.d. *de los menores $i \times i$ de A, entonces $\Delta_i = d_1 \cdots d_i$, para $1 \leq i \leq r$.*

Demostración: *a*) Definimos la **longitud** de a, escrita $\ell(a)$, como el número de primos (con sus multiplicidades) que aparecen en la factorización de a, de manera que $\ell(a) = 0 \Leftrightarrow a \in k^\star$.

Dada $A \in M_{m\times n}(k)$, elegimos una matriz $B = (b_{ij})$ con una coordenada b_{pq} tal que $\ell(b_{pq})$ sea mínimo entre las coordenadas de matrices equivalentes con A. Podemos suponer que $p = q = 1$ para tener $\ell(b_{11})$ mínimo.

Afirmamos que $b_{11} \mid b_{1j}$ y que $b_{11} \mid b_{i1}$ para todos $i, j > 1$. Veamos qué pasaría si $b_{11} \nmid b_{12}$: Sea $d = (b_{11}, b_{12})$, de manera que $\ell(d) < \ell(b_{11})$ y existirían $x, y \in k$ con $d = xb_{11} + yb_{12}$.

Sean $u = b_{12}/d,\ v = -b_{11}/d \in k$, por lo que

$$\begin{pmatrix} -v & u \\ y & -x \end{pmatrix} \begin{pmatrix} x & u \\ y & v \end{pmatrix} = 1.$$

Así, tendríamos que

$$B \begin{pmatrix} x & u & & & \\ y & v & & & \\ & & 1 & & \mathbf{0} \\ & & & \ddots & \\ & & \mathbf{0} & & 1 \end{pmatrix} = \begin{pmatrix} d & 0 & * & \cdots & * \\ & & & & \\ & & & \mathbf{*} & \\ & & & & \end{pmatrix}$$

con $\ell(d) < \ell(b_{11})$, que es una contradicción.

Establecida la afirmación, podemos suponer que $b_{1j} = b_{i1} = 0$ para todos $i, j > 1$. Entonces tenemos que $b_{11} \mid b_{ij}$ para todos i, j; pues de lo contrario, al sumar el renglón i al renglón 1 contradiríamos al paso anterior. Concluimos por inducción en n ó en m, ya que la submatriz

$$\begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

se puede "diagonalizar" preservándose la condición $b_{11} \mid b_{ij}$.

b) Si multiplicamos a la matriz A por la izquierda o por la derecha por una matriz cuadrada invertible P del tamaño adecuado, tendremos que todo menor $i \times i$ del resultado estará en el ideal (Δ_i). Al ser P invertible, esto implica que el m.c.d. de los menores $i \times i$ es el mismo en ambos casos. Así, los Δ_i son invariantes ante equivalencia. Cuando A es diagonal como en *a*), con $d_i \mid d_j$ para todo $i < j$, se tienen las igualdades $\Delta_i = d_1 \cdots d_i$, para todo $1 \leq i \leq r$, de donde se obtiene la unicidad de los factores d_i. □

Decimos que una matriz "diagonal" como en *a*) está en **forma canónica** y que los números d_i son los **factores invariantes** de A. Observemos que r es el rango determinantal de A y que también es un invariante de A ante equivalencia.

Existen paquetes de computación como Macaulay2, CoCoA y Singular, que efectúan el cálculo directo de los ideales Δ_i, por lo que no es necesario realizar operaciones elementales para encontrar la forma canónica de una matriz dada. Sin embargo, en muchos casos es deseable encontrar las matrices P ó Q que llevan una matriz dada a su forma canónica.

Ejemplo. Podemos llevar la matriz $A \in M_3(\mathbb{Z})$ a su forma canónica B:

$$A = \begin{pmatrix} 12 & -12 & 240 \\ 4 & 9 & 81 \\ 0 & 480 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 480 \end{pmatrix} = B,$$

pues $\Delta_1 = 1$, $\Delta_2 = 12$ y $\Delta_3 = (12)(480)$.

Ejercicios

1. Encuentre un conjunto completo de invariantes ante equivalencia para matrices $m \times n$ sobre un campo.
2. Sean k un dominio principal y $d, a_1, ..., a_n \in k$ tales que valga la igualdad de ideales $(d) = (a_1, ..., a_n)$. Demuestre que existe $A \in GL_n(k)$ tal que $(a_1, a_2, ..., a_n)A = (d, 0, ..., 0)$.
3. Sean k un dominio principal y $A, B \in M_n(k)$ con $\det AB \neq 0$. Si $\operatorname{diag}(a_1, ..., a_n)$, $\operatorname{diag}(b_1, ..., b_n)$ y $\operatorname{diag}(c_1, ..., c_n)$ son formas canónicas para A, B y $C = AB$ respectivamente, demuestre que $a_i \mid c_i$ y $b_i \mid c_i$ para todo $1 \leq i \leq n$. (Sugerencias: Considere el caso en que A y B son diagonales, suponga que k contiene un único elemento primo).

4.5 Módulos sobre Dominios Principales

En esta sección, suponemos que nuestros módulos son finitamente generados sobre un dominio principal k. El módulo generado por $\{a_1, ..., a_m\}$ lo escribimos así: $(a_1, ..., a_m)$, o bien así: $ka_1 + \cdots + ka_m$.

Teorema 4.24 *a) Todo submódulo N de un módulo libre M, es libre con* $\operatorname{rango} N \leq \operatorname{rango} M$.

b) Existen bases $\{u_1, ..., u_m\}$ de M y $\{v_1, ..., v_n\}$ de N tales que $v_i = a_i u_i$ para $1 \leq i \leq n$, con $a_1 \mid a_2 \mid \cdots \mid a_n$.

Demostración: *a)* Procedemos por inducción en $m = \operatorname{rango} M$, siendo claro el resultado para $m = 1$. Sea $\{u_1, ..., u_m\}$ una base de M.

Si $N \subseteq (a_1, ..., a_{m-1})$, entonces concluimos por la hipótesis inductiva. Si no, existe $u = c_1u_1 + \cdots + c_mu_m \in N$ con $c_m \neq 0$. El conjunto de tales coeficientes c_m forma un ideal (b) de k, para el que existe

$$w = c'_1u_1 + \cdots + c'_{m-1}u_{m-1} + bu_m \in N.$$

La hipótesis inductiva nos permite encontrar una base $\{v_1, ..., v_{n-1}\}$ para $N \cap (u_1, ..., u_{m-1})$, con $n \leq m$; pero entonces $\mathrm{A} = \{v_1, ..., v_n\}$, con $v_n = w$, es una base de N: Claramente, A genera a N; y es fácil ver que A es linealmente independiente.

b) Si $\Psi : N \hookrightarrow M$ es la inclusión, entonces existen bases de M y de N tales que con respecto a ellas, la matriz de Ψ está en la forma canónica del Teorema 4.23. □

Se dice que un k-módulo M es **cíclico** cuando admite a un conjunto $\{a\}$ con un único elemento como generador. El **anulador de un elemento** $a \in M$ es el ideal de k, escrito $\mathrm{an}(a)$, definido como $\{r \in k \mid ra = 0\}$. El **anulador de un módulo** M es el ideal $\{r \in k \mid rM = 0\}$, escrito $\mathrm{an}(M)$.

Un k-módulo cíclico $M = (a)$ queda completamente descrito por el anulador de su generador, pues $M \cong k/\mathrm{an}(a)$. Observemos que si $b \in (a)$, entonces $\mathrm{an}(a) \subseteq \mathrm{an}(b)$, de manera que si b es otro generador de (a), entonces $\mathrm{an}(a) = \mathrm{an}(b)$. En particular, si $\mathrm{an}(a) = 0$, tendremos el isomorfismo de k-módulos $(a) \cong k$.

Teorema 4.25 *Sea $M \neq (0)$ un k-módulo finitamente generado. Entonces M es una suma directa de módulo cíclicos:*

$$M = ka_1 \oplus \cdots \oplus ka_s, \text{ donde } (d_1) \supseteq \cdots \supseteq (d_s), \ \mathrm{an}(a_i) = (d_i) \neq k, \forall\ i.$$

Demostración: Como M es finitamente generado, existen un k-módulo libre $k^{(n)}$ y un morfismo suprayectivo $\varphi : k^{(n)} \to M$. Sean $\ker \varphi \cong k^{(m)}$ y $\psi : k^{(m)} \to k^{(n)}$ la inclusión.

Elegimos bases adecuadamente, para tener a $\{u_1, ..., u_n\}$ como base de $k^{(n)}$ y como matriz asociada a ψ a una de forma $\mathrm{diag}(d'_1, ..., d'_r, 0, ...)$ con $d'_i \neq 0$ para todo i, con $d'_i \mid d'_{i+1}$ para $i = 1, ..., r-1$.

Si $d'_1, ..., d'_t = 1$ (ó bien son unidades), vemos que M está generado por las imágenes $a_1 = \varphi(u_{t+1}), ..., a_s = \varphi(u_n)$, al escribir $s = n - t$. Además, si $d_1 = d'_{t+1}, ..., d_s = d'_n$, tendremos que $M = ka_1 \oplus \cdots \oplus ka_s$, con ideales $\mathrm{an}(a_i) = d_i$ como en el enunciado. □

Al **submódulo de torsión** de M, escrito $\mathrm{tor}\, M$, lo definimos como $\mathrm{tor}\, M = \{a \in M \mid \text{existe } 0 \neq r \in k \text{ con } ra = 0\}$. Se dice que M es **de torsión** cuando $M = \mathrm{tor}\, M$; o que M es **libre de torsión** cuando $\mathrm{tor}\, M = 0$.

Teorema 4.26 *Todo módulo finitamente generado sobre un dominio principal, es la suma directa de su submódulo de torsión y de un módulo libre.*

Demostración: Si $M = ka_1 \oplus \cdots \oplus ka_s$ con $\mathrm{an}(a_1) \supseteq \cdots \supseteq \mathrm{an}(a_s)$, donde $\mathrm{an}(a_r) \neq 0$; pero $\mathrm{an}(a_{r+1}) = 0$, entonces $ka_1 \oplus \cdots \oplus ka_r \subseteq \mathrm{tor}\, M$.

Recíprocamente, si $m = c_1a_1 + \cdots + c_sa_s \in \mathrm{tor}\, M$, entonces existe $0 \neq c \in k$ tal que $0 = cm = cc_1a_1 + \cdots + cc_sa_s$, por lo que para $i > r$ se tiene que $cc_i = 0$ y que $c_i = 0$. Así, $ka_1 \oplus \cdots \oplus ka_r = \mathrm{tor}\, M$.

Por otra parte, $ka_{r+1} \oplus \cdots \oplus ka_s \cong k^{(s-r)}$ es libre, por lo que finalmente $M = (\operatorname{tor} M) \oplus k^{(s-r)}$. □

Sabemos que todo dominio principal es de factorización única. Dado un elemento primo $p \in k$, definimos la **p-componente primaria** M_p de un k-módulo M como el conjunto $\{z \in M \mid p^i z = 0 \text{ para algún } i \in \mathbb{N}\}$. Es inmediato que M_p es un submódulo de $\operatorname{tor} M$; y que si $p_1, ..., p_r$ son distintos primos, entonces $M_{p_1}, ..., M_{p_r}$ forman suma directa.

Lema 4.27 *Sea $M = ka$ un módulo cíclico con $\operatorname{an}(a) = (r)$, donde $r = r_1 r_2$ para r_1 y r_2 elementos primos relativos en k; entonces $M = kb \oplus kc$, con $\operatorname{an}(b) = (r_1)$ y $\operatorname{an}(c) = (r_2)$. Recíprocamente, si $M = kb \oplus kc$, con $\operatorname{an}(b) = (r_1)$ y $\operatorname{an}(c) = (r_2)$, donde r_1 y r_2 son primos relativos; entonces $M = ka$ con $\operatorname{an}(a) = (r_1 r_2)$.*

Demostración: A partir de a, definimos $b = r_2 a$ y $c = r_1 a$, para tener que $\operatorname{an}(b) = (r_1)$ y que $\operatorname{an}(c) = (r_2)$. Además, existen $s, t \in k$ tales que $1 = sr_1 + tr_2$, por lo que $a = 1a = (sr_1 + tr_2)a \in kb + kc$; pero $kb \cap kc = 0$, pues $m \in kb \cap kc \Rightarrow r_1 m = 0 = r_2 m \Rightarrow 1m = (sr_1 + tr_2)m = 0$. Así, $M = kb \oplus kc$.

Recíprocamente, dado que $M = kb \oplus kc$, definimos $a = b + c$. Aquí, $ra = 0$ implica $rb = 0 = rc$, porque $kb \cap kc = 0$; y entonces $r \in (r_1 r_2)$. Como $r_1 r_2 a = 0$, vemos que $\operatorname{an}(a) = (r_1 r_2)$. Ahora bien, existen $s, t \in k$ con $1 = sr_1 + tr_2$, por lo que $b = 1b = (sr_1 + tr_2)b = tr_2 b = tr_2(b + c) = tr_2 a \in ka$. Análogamente, $c \in ka$; y así $ka = M$. □

Teorema 4.28 *Si M es un módulo de torsión finitamente generado, entonces $M_p = 0$ para todo primo $p \in k$, con un número finito de excepciones: $p_1, ..., p_r$; y entonces $M = M_{p_1} \oplus \cdots \oplus M_{p_r}$. También existe una descomposición $M = ka_1 \oplus \cdots \oplus ka_s$, donde cada $\operatorname{an}(a_i)$ es de forma (p^t), con p primo y $t \geq 1$.*

Demostración: Si $\{b_1, ..., b_n\}$ genera a M con $\operatorname{an}(b_i) = (s_i)$ para cada i, hacemos la lista $p_1, ..., p_r$ de los primos en k que dividen a los s_i, para tener que cada $kb_i \subseteq M_{p_1} \oplus \cdots \oplus M_{p_r}$; y una descomposición de M como suma directa de módulos cíclicos primarios, gracias al lema.

Si q es un primo distinto de todo p_i, entonces

$$M_q = M_q \cap M = M_q \cap (M_{p_1} \oplus \cdots \oplus M_{p_r}) = 0.$$

Por el Lema 4.27, cada submódulo cíclico ka_i ó kb_j puede descomponerse como suma directa de módulos cíclicos primarios, es decir, con anuladores de forma (p^t) con p primo. □

Para M finitamente generado, tenemos $M = (\operatorname{tor} M) \oplus M'$, con M' libre. El **rango** de M es $\operatorname{rango}(M/\operatorname{tor} M)$, que no depende del módulo M'.

Teorema 4.29 *a) Si tenemos* $M = ka_1 \oplus \cdots \oplus ka_r = kb_1 \oplus \cdots \oplus kb_s$, *donde* $k \neq \text{an}(a_1) \supseteq \cdots \supseteq \text{an}(a_r)$ *y también* $k \neq \text{an}(b_1) \supseteq \cdots \supseteq \text{an}(b_s)$. *Entonces* $r = s$ *y también* $\text{an}(a_i) = \text{an}(b_i)$, *para* $1 \leq i \leq r$.

b) Si M *es de torsión y* $M = kc_1 \oplus \cdots \oplus kc_r = kd_1 \oplus \cdots \oplus kd_s$ *con cada* $\text{an}(c_i)$ *y cada* $\text{an}(d_j)$ *de forma* (p^t), *entonces* $r = s$ *y* $\text{an}(c_i) = \text{an}(d_i)$, *para* $1 \leq i \leq r$.

Demostración: Las hipótesis $\text{an}(a_1) \neq k \neq \text{an}(b_1)$ nos garantizan que $ka_i \neq 0 \neq kb_j$, para todos i, j. El número de índices i tales que $\text{an}(a_i) = 0$ es el rango de $(M/\operatorname{tor} M)$; y coincide con el número de índices j tales que $\text{an}(b_j) = 0$, por la observación previa. Esto nos permite suponer que $M = \operatorname{tor} M$.

Los anuladores de los elementos a_i ó b_j pueden recuperarse a partir de los anuladores primarios así: $\text{an}(a_1)$ es el m.c.m. de todos ellos, $\text{an}(a_2)$ es el m.c.m. de los anuladores primarios que queden al eliminar aquellos cuyo producto es $\text{an}(a_1)$, etc.

Así, la unicidad de los anuladores primarios implica la unicidad de nuestros $\text{an}(a_i)$. Por tanto, ahora suponemos que $M = M_p$ es primario. Escribamos pues $\text{an}(c_i) = (p^{\alpha_i})$ y $\text{an}(d_j) = (p^{\beta_j})$, para tener

$$\alpha_1 \leq \cdots \leq \alpha_r, \ \beta_1 \leq \cdots \leq \beta_s.$$

Para cada $t \in \mathbb{N}$ definimos $p^t M = \{p^t m \mid m \in M\}$, un submódulo de M. También definimos $M_t = p^t M / p^{t+1} M$, que es un espacio vectorial sobre $k/(p)$ de manera natural, y de dimensión finita, independiente de cualquier descomposición de M.

Observando que $\dim M_t$ es el número de sumandos cíclicos primarios kz con $\text{an}(z) = (p^w)$, donde $w > t$, concluimos que los números α_i (ó β_j) quedan determinados por esas dimensiones. Así, $r = s$ y también $\text{an}(c_i) = \text{an}(d_i)$, para $1 \leq i \leq r$. □

Los ideales $\text{an}(a_1), ..., \text{an}(a_r)$ se llaman **ideales factores invariantes** de M. Los ideales $\text{an}(z) = (p^w)$, con p primo, de una descomposición en sumandos cíclicos primarios de un módulo de torsión, se llaman **ideales divisores elementales**.

Cuando $k = \mathbb{Z}$, los ideales anteriores admiten generadores positivos; y cuando $k = K[T]$, con K un campo y T una variable, admiten generadores mónicos. Estos nuevos generadores se llaman **factores invariantes** ó **divisores elementales**, respectivamente.

Aplicación a Grupos Abelianos

Cuando $k = \mathbb{Z}$, un k-módulo es lo mismo que un grupo Abeliano. Decimos que un conjunto dado de invariantes ante cierta relación de equivalencia es **completo** cuando dos objetos están en la misma clase de equivalencia si y sólo si tienen iguales los invariantes del conjunto dado.

El siguiente teorema, que es inmediato, sintetiza todo lo obtenido en la Sección 1.13.

Teorema 4.30 *a) Todo grupo Abeliano finitamente generado es la suma directa de un grupo finito único (su torsión) y de un grupo libre Abeliano. El rango de la componente libre es un invariante.*

b) Todo grupo Abeliano finito es la suma directa de grupos cíclicos de órdenes potencias de primos. Estos órdenes con sus multiplicidades son únicos y constituyen un conjunto completo de invariantes del grupo.

c) Todo grupo Abeliano finito es la suma directa de grupos cíclicos de órdenes $d_1, ..., d_n$, donde $d_i \mid d_{i+1}$ para $1 \leq i < n$. Estos órdenes con sus multiplicidades son únicos y constituyen un conjunto completo de invariantes del grupo.

Ejercicios

Aquí, Z_n es el grupo cíclico de orden n.

1. Dados un número primo p y un entero positivo n, demuestre que Z_{p^n} no es la suma directa de dos subgrupos propios.

2. Demuestre que el ideal $(2, X)$ de $\mathbb{Z}[X]$ no es suma directa de dos o más $\mathbb{Z}[X]$-módulos cíclicos no triviales.

3. Sean k un dominio principal y M un k-módulo finitamente generado de torsión, con ideales factores invariantes $(d_1) \supseteq \cdots \supseteq (d_n)$. Demuestre que todo submódulo y toda imagen homomorfa de M son de torsión con factores invariantes $(e_1) \supseteq \cdots \supseteq (e_s)$ tales que $s \leq n$; y además $e_s \mid d_n,\ e_{s-1} \mid d_{n-1}, ...,\ e_1 \mid d_{n-s+1}$.

4. Encuentre a los grupos Abelianos G no isomorfos entre sí, tales que exista una sucesión exacta $0 \to Z_4 \to G \to Z_{16} \to 0$.

5. Sea $G = \langle a_1, a_2, a_3 \mid 30a_1 + 10a_2 + 16a_3 = 0, 4a_1 + 2a_2 + 2a_3 = 0, 24a_1 + 8a_2 + 14a_3 = 0\rangle$. Expréselo como suma directa de grupos cíclicos de órdenes potencias de primos.

4.6 Similaridad de Matrices sobre Campos

Sean k un campo, V un espacio vectorial de dimensión finita y $T : V \to V$ una transformación lineal. Al elegir una base $\{u_1, ..., u_n\}$ de V, obtenemos la matriz $A = (a_{ij}) \in M_n(k)$ como sigue:

$$T(u_i) = \sum_{j=1}^{n} a_{ij}u_j, \text{ para } 1 \leq i \leq n.$$

Si elegimos una nueva base $\{v_1, ..., v_n\}$ de V relacionada con la base original por medio de la matriz $P = (p_{ij}) \in M_n(k)$, donde

$$v_i = \sum_{j=1}^{n} p_{ij} u_j, \text{ para } 1 \leq i \leq n;$$

la Proposición 4.2 garantiza que P es invertible y el Teorema 4.3 implica que la matriz que le corresponde a T con respecto a la nueva base es PAP^{-1}.

Decimos que dos matrices $A, B \in M_n(k)$ son **similares** cuando existe $P \in M_n(k)$ invertible tal que $B = PAP^{-1}$. Claramente, similaridad es una relación de equivalencia.

Supongamos fija la transformación T, y consideremos el anillo de polinomios $R = k[X]$ en la variable X con coeficientes en k. Hacemos de V un R-módulo declarando que la variable X actúa precisamente como la transformación lineal T.

Nos proponemos utilizar los resultados de la sección anterior para entender al R-módulo V, el cual es finitamente generado como R-módulo por cualquier base del espacio vectorial. Además, V es de torsión, pues dado $0 \neq u \in V$, tendremos que el conjunto $\{u, T(u), T^2(u), ...\}$ será linealmente dependiente, por lo que un polinomio no trivial en T anulará a u.

Proposición 4.31 *Se tiene la sucesión exacta* $R^{(n)} \xrightarrow{\varphi} R^{(n)} \xrightarrow{\psi} V \to 0$, *donde* $\psi(e_i) = u_i$ *para todo* $1 \leq i \leq n$; *y donde* φ *tiene como matriz a*

$$X - A = \begin{pmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & X - a_{nn} \end{pmatrix},$$

con respecto a $\{e_1, ..., e_n\}$, *base natural de* $R^{(n)}$.

Demostración: Claramente, ψ es suprayectivo. Como

$$\psi(Xe_i - \sum_{j=1}^{n} a_{ij} e_j) = Tu_i - \sum_{j=1}^{n} a_{ij} u_j = 0, \text{ para todo } 1 \leq i \leq n,$$

se tiene que $\operatorname{Im} \varphi \subseteq \ker \psi$.

Para ver la inclusión recíproca, supongamos que $N = \operatorname{Im} \varphi$; y que $x = \sum_{i=1}^{n} p_i(X) e_i \in \ker \psi$, con cada $p_i(X) \in R$. Debido a que para todo i vale $Xe_i - \sum_{j=1}^{n} a_{ij} e_j \in N$, tenemos que existen $c_j \in k$ tales que

$$x \equiv \sum_{j=1}^{n} c_j e_j \pmod{N};$$

pero entonces

$$\psi(x) = \sum_{j=1}^{n} c_j u_j = 0 \Rightarrow c_j = 0, \ \forall j.$$

Así, $x \in N$; y la sucesión es exacta. □

Cuando se tiene una sucesión es exacta $M \xrightarrow{f} N \xrightarrow{g} P \to 0$, se dice que P es el **conúcleo** de f. Así, $V = \operatorname{coker} \varphi = \operatorname{coker}(X - A)$.

La sucesión exacta $R^{(n)} \xrightarrow{\varphi} R^{(n)} \xrightarrow{\psi} V \to 0$, ó bien la matriz $X - A$ constituyen una presentación del R-módulo V.

Gracias al Teorema 4.23, sabemos que al morfismo φ le corresponde una matriz de forma $\operatorname{diag}(1, ..., 1, d_1, ..., d_r)$ con $d_1 \mid d_2 \mid \cdots \mid d_r$, donde $\operatorname{gr} d_i \geq 1$ para todo i; y donde $d_1 \cdots d_r = \det P(X - A)Q$ es asociado en R del **polinomio característico** de A:

$$\det(X - A) = X^n + a_{n-1}X^{n-1} + \cdots + a_0,$$

donde $-a_{n-1} = \operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$ y donde $(-1)^n a_0 = \det A$.

Corolario 4.32 *Si $P(X - A)Q = \operatorname{diag}(1, ..., 1, d_1, ..., d_r)$ con P, Q invertibles, $(d_1) \neq R$ y $Q^{-1} = (q_{ij})$, entonces $V = Rv_{n-r+1} \oplus \cdots \oplus Rv_n$, donde*

$$v_i = \sum_{j=1}^{n} q_{ij}u_j \quad y \quad \operatorname{an}(v_{n-r+1}) = (d_1) \supseteq \cdots \supseteq (d_r) = \operatorname{an}(v_n).$$

Demostración: El Teorema 4.3 nos da $V = Rv_1 \oplus \cdots \oplus Rv_n$, para

$$v_i = \sum_{j=1}^{n} q_{ij}u_j, \text{ donde } \operatorname{an}(v_1) = \cdots = \operatorname{an}(v_{n-r}) = R,$$

mientras que $R \neq an\ (v_{n-r+1}) = (d_1) \supseteq \cdots \supseteq (d_r) = \operatorname{an}(v_n)$. □

Sea $J = \{p(X) \in R \mid p(T) = 0\}$, un ideal propio de R, pues al ser V un R-módulo de torsión, existen $0 \neq p_i(X) \in R$ tales que $p_i(T)u_i = 0$ para $1 \leq i \leq n$; y entonces $p_1(X) \cdots p_n(X) \in J$. Si $J = (m(X))$ con $m(X)$ mónico, decimos que $m(X)$ es el **polinomio mínimo** de T.

Teorema 4.33 (Cayley-Hamilton-Frobenius) *Con esta notación,*

$$\det(X - A) = m(X)\Delta_{n-1},$$

donde Δ_{n-1} es el m.c.d. *en R de los menores $(n-1) \times (n-1)$ de la matriz $X - A$. En particular, T es raíz del polinomio* $\det(X - A) = 0$.

Demostración: En el corolario se ve que $(m(X)) = (d_r)$; pero sabemos que $\Delta_n = d_r\Delta_{n-1}$, con $\Delta_n = \det(X - A)$, por el Teorema 4.23. □

Cuando $V = Rv$, decimos que V es **cíclico**, o bien que v es **cíclico**.

Teorema 4.34 (Forma Canónica Racional) *Sean k un campo arbitrario y $A \in M_n(k)$ una matriz cuadrada. Entonces existe una matriz invertible $P \in M_n(k)$ tal que PAP^{-1} es de forma*

$$\begin{pmatrix} C_1 & & & \\ & C_2 & & \mathbf{0} \\ & & \ddots & \\ & \mathbf{0} & & C_r \end{pmatrix},$$

donde cada bloque C_i es la ***matriz compañera***

$$C_i = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{m-1} \end{pmatrix}$$

de un polinomio $d_i(X) = X^m + c_{m-1}X^{m-1} + \cdots + c_0$. Esta expresión es única si $d_1(X) \mid d_2(X) \mid \cdots \mid d_r(X)$ y cada $d_i(X)$ es mónico. Alternativamente, la expresión es única hasta reacomodo de los bloques centrales si cada $d_i(X)$ es mónico y es potencia de un polinomio irreducible.

Demostración: Es suficiente considerar el caso cíclico Rv con $\mathrm{an}(v) = X^m + c_{m-1}X^{m-1} + \cdots + c_0$. Aquí, Rv admite como base al conjunto $\{v, Xv, X^2v, ..., X^{m-1}v\}$, donde la acción de T es como sigue:

$$\begin{array}{rcll} Tv & = & Xv & \\ TXv & = & X^2v & \\ & \vdots & & \\ TX^{m-1}v & = & X^m v & = \quad -c_0v - c_1Xv - \cdots - c_{m-1}X^{m-1}v \end{array}$$

Así, V es suma directa de subespacios cíclicos, en cada uno de los cuales T tiene asociada una matriz C_i con respecto a una base adecuada. □

Ejemplo. Encontraremos $P \in M_3(\mathbb{Q})$ invertible tal que PAP^{-1} esté en su forma canónica racional para

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \in M_3(\mathbb{Q}).$$

$$X - A = \begin{pmatrix} X & 0 & 1 \\ 0 & X-1 & 0 \\ -1 & 0 & X+1 \end{pmatrix} \xrightarrow{\mathrm{r}} \begin{pmatrix} 0 & 0 & X^2+X+1 \\ 0 & X-1 & 0 \\ -1 & 0 & X+1 \end{pmatrix}$$
$$\xrightarrow{\mathrm{r}} \begin{pmatrix} 1 & 0 & -X-1 \\ 0 & X-1 & 0 \\ 0 & 0 & X^2+X+1 \end{pmatrix} \xrightarrow{\mathrm{c}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & X^2+X+1 \end{pmatrix}$$

Así, tenemos que existen $B, C \in M_3(\mathbb{Q}[X])$ invertibles tales que

$$B(X-A)C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & X^2+X+1 \end{pmatrix}.$$

Como solamente hubo una operación columna, vemos que

$$C = \begin{pmatrix} 1 & 0 & X+1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ y que } C^{-1} = \begin{pmatrix} 1 & 0 & -X-1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La forma canónica racional de A es:

$$\left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & -1 & -1 \end{array}\right).$$

Si la base original era $\{u_1, u_2, u_3\}$, la nueva base se obtiene así:

$$\begin{array}{lcccl} v_0 & = & u_1 - (X+1)u_3 & = & 0 \\ v_1 & = & u_2 & & \\ v_2 & = & u_3 & & \\ v_3 & = & Xu_3 & = & u_1 - u_3 \end{array}$$

Matriz que produce el cambio de base:

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

Se puede verificar directamente que:

$$PAP^{-1} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & -1 & -1 \end{array}\right).$$

Teorema 4.35 (Forma Canónica de Jordan) *Sean k un campo algebraicamente cerrado y $A \in M_n(k)$ una matriz cuadrada. Entonces existe una matriz invertible $P \in M_n(k)$ tal que PAP^{-1} es de forma*

$$\begin{pmatrix} J_1 & & & \mathbf{0} \\ & J_2 & & \\ & & \ddots & \\ \mathbf{0} & & & J_r \end{pmatrix},$$

donde cada bloque J_i es de ***Jordan***

$$J_i = \begin{pmatrix} \lambda & 1 & & \mathbf{0} & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & \mathbf{0} & & & \lambda \end{pmatrix}$$

Esta expresión es única hasta reacomodo de los bloques de Jordan.

Demostración: Es suficiente ver el caso cíclico Rv con $\mathrm{an}(v) = (X-\lambda)^m$, con $\lambda \in k$. Aquí, el conjunto $\{v, (X-\lambda)v, (X-\lambda)^2 v, ..., (X-\lambda)^{m-1}v\}$ es una base de Rv; y la acción de T es como sigue:

$$\begin{array}{rcrcl} Tv & = & Xv & = & \lambda v + (X-\lambda)v \\ T(X-\lambda)v & = & X(X-\lambda)v & = & \lambda(X-\lambda)v + (X-\lambda)^2 v \\ & & \vdots & & \vdots \\ T(X-\lambda)^{m-1}v & = & X(X-\lambda)^{m-1}v & = & \lambda(X-\lambda)^{m-1}v \\ & & & & +(X-\lambda)^m v \\ & & & = & \lambda(X-\lambda)^{m-1}v \end{array}$$

Así, V es suma directa de subespacios cíclicos, en cada uno de los cuales T tiene asociada la matriz J_i con respecto a una base adecuada. □

Ejemplo. Buscamos $P \in M_3(\mathbb{C})$ invertible tal que PAP^{-1} esté en su forma canónica de Jordan para

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \in M_3(\mathbb{C}).$$

En el ejemplo anterior, partimos de la base $\{u_1, u_2, u_3\}$ y encontramos que $V = Rv_1 \oplus Rv_2$, con $v_1 = u_2$ y $\mathrm{an}(v_1) = (X-1)$ y con $v_2 = u_3$ satisfaciendo $\mathrm{an}(v_2) = (X^2 + X - 1)$.

Sea ω es una raíz cúbica primitiva de la unidad. Tenemos que

$$X^2 + X - 1 = (X-\omega)(X-\omega^2).$$

Como $\mathrm{an}[(X-\omega^2)v_2] = (X-\omega)$ y como $\mathrm{an}[(X-\omega)v_2] = (X-\omega^2)$, podemos considerar la base de V:

$$\{u_2, (X-\omega^2)v_2 = u_1 - u_3 - \omega^2 u_3, (X-\omega)v_2 = u_1 - u_3 - \omega u_3\}$$

para tener que

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1-\omega^2 \\ 1 & 0 & -1-\omega \end{pmatrix} \Rightarrow PAP^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}.$$

Teorema 4.36 (Forma Canónica Real) *Sea $A \in M_n(\mathbb{R})$ una matriz cuadrada. Entonces existe una matriz invertible $P \in M_n(\mathbb{R})$ tal que PAP^{-1} es de forma*

$$\begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_r \end{pmatrix},$$

*donde cada bloque A_i es de **Jordan** $J_i(\lambda)$ como en el teorema anterior con $\lambda \in \mathbb{R}$ ó bien es de forma*

$$\left(\begin{array}{cc|cc|cc|cc} 0 & 1 & 0 & 0 & & & & \\ -b & -a & 1 & 0 & & & & \\ \hline & & 0 & 1 & 0 & 0 & & \\ & & -b & -a & 1 & 0 & & \\ \hline & & & & \ddots & \ddots & \ddots & \\ & & & & & \ddots & \ddots & \ddots \\ \hline & & & & & & 0 & 1 & 0 & 0 \\ & & & & & & -b & -a & 1 & 0 \\ \hline & & & & & & & & 0 & 1 \\ & & & & & & & & -b & -a \end{array}\right),$$

con $a^2 < 4b$. Esta expresión es única hasta reacomodo de los bloques A_i.

Demostración: Los polinomios irreducibles de $\mathbb{R}[X]$ son de forma $X - \lambda$ ó bien $X^2 + aX + b$ con $a^2 - 4b < 0$. Conocemos el caso cíclico Rv, donde $\text{an}(v) = (X - \lambda)^m$ con $\lambda \in \mathbb{R}$. Para el último caso, supongamos que Rv satisface $\text{an}(v) = (X^2 + aX + b)^m$ con $a^2 < 4b$. Aquí, el conjunto

$$\{v, Xv, (X^2 + aX + b)v, X(X^2 + aX + b)v, (X^2 + aX + b)^2 v, ... \\ ..., (X^2 + aX + b)^{m-1}v, X(X^2 + aX + b)^{m-1}v\}$$

es una base de Rv; y la acción de T es como sigue:

$$\begin{aligned} Tv &= Xv \\ T(Xv) = X^2v &= -bv - a(Xv) + (X^2 + aX + b)v \\ T[(X^2 + aX + b)v] &= X(X^2 + aX + b)v \\ &\vdots \\ T[(X^2 + aX + b)^{m-1}v] &= X(X^2 + aX + b)^{m-1}v \\ T[X(X^2 + aX + b)^{m-1}v] &= X^2(X^2 + aX + b)^{m-1}v \\ = -b(X^2 + aX + b)^{m-1}v &- aX(X^2 + aX + b)^{m-1}v \end{aligned}$$

Así, V es suma directa de subespacios cíclicos, en cada uno de los cuales T tiene asociada una matriz A_i con respecto a una base adecuada. □

Ejercicios

1. *a*) Demuestre que $0 \to R^{(n)} \xrightarrow{\varphi} R^{(n)} \xrightarrow{\psi} V \to 0$, con φ y ψ como en la Proposición 4.31, es una sucesión exacta.

   b) Demuestre que $\{Xe_i - \sum_{j=1}^{n} a_{ij}e_j \mid 1 \leq i \leq n\}$ es una base para $\ker \psi$.

2. Si k es un campo y $A, B \in M_n(k)$, demuestre que $X - A$ y $X - B$ son equivalentes en $M_n(k[X])$ si y sólo si A y B son similares en $M_n(k)$.

3. Sean $A, B \in M_n(\mathbb{C})$. Demuestre que A y B son similares en $M_n(\mathbb{C})$ si y sólo si $(c - A)^m$ y $(c - B)^m$ tienen el mismo rango para todos $c \in \mathbb{C}$ y $m \in \mathbb{N}$.

4. Sean k un campo y $A \in M_n(k)$ tal que $A^2 = A$. Demuestre que A es similar con

$$\left(\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ \hline & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{array}\right),$$

5. Demuestre que toda matriz $A \in M_n(k)$, con k un campo, es similar con su transpuesta.

6. Si p es primo, demuestre que $A, B \in M_p(\mathbb{Z}/p\mathbb{Z})$ son similares si

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix} \text{ y } B = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}.$$

7. Encuentre el polinomio mínimo de

$$A = \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 19 & -23 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 19 & -23 \end{array}\right).$$

8. Demuestre que si $A \in M_2(\mathbb{R})$ y $A^2 = -1$, entonces A es similar con

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

4.7 La Descomposición de Jordan-Chevalley

En esta sección continuamos estudiando la estructura de un espacio vectorial con respecto a su interacción con una transformación lineal dada. Los métodos aquí usados no involucran la estructura de módulos sobre dominios principales. Hasta nuevo aviso, suponemos lo siguiente: V es un espacio vectorial de dimensión n sobre un campo algebraicamente cerrado k; y $T : V \to V$ es una transformación lineal.

Se dice que T es **diagonalizable** o **semisimple** cuando existe una base de V con respecto a la cual a T le corresponde una matriz diagonal.

Decimos que $\lambda \in k$ y $v \in V$ son respectivamente un **valor característico** y un **vector característico** de T cuando $T(v) = \lambda v$.

Teorema 4.37 *Las siguientes condiciones en T son equivalentes:*
a) T es diagonalizable.
b) V tiene una base que consiste de vectores característicos de T.
c) V es suma directa de subespacios en los que T actúa como escalar.
d) T satisface algún polinomio en $k[X]$ con raíces distintas.
e) El polinomio mínimo de T es separable.

Demostración: Las equivalencias *a)* $\Leftrightarrow$ *b)* $\Leftrightarrow$ *c)* y *d)* $\Leftrightarrow$ *e)* son inmediatas.

Veamos que *a)* $\Rightarrow$ *d)*. Si T está asociada a la matriz

$$\left(\begin{array}{ccc|c|cccc} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ \hline & & & \ddots & & & & \\ \hline & & & & \lambda_s & & & \\ & & & & & \ddots & & \\ & & & & & & \lambda_s & \end{array}\right),$$

con respecto a alguna base y $\lambda_i \neq \lambda_j$ siempre que $i \neq j$, entonces $f(X) = (X - \lambda_1) \cdots (X - \lambda_s)$ es separable y $f(T) = 0$.

d) $\Rightarrow$ *c)*: Si $f(X) = (X - a_1) \cdots (X - a_s)$ con $a_i \neq a_j$ siempre que $i \neq j$ y $f(T) = 0$, escribimos $f_i(X) = f(X)/(X - a_i) \in k[X]$ para $1 \leq i \leq s$. Es inmediato que se tiene la igualdad de "ideales" $k[X] = (f_1(X), ..., f_s(X))$, por lo que existen polinomios $g_1, ..., g_s$ tales que $g_1 f_1 + \cdots + g_s f_s = 1$.

Los subespacios $V_i = \ker(T - a_i)$ forman suma directa. Dado $v \in V$, podemos escribir $v = [g_1(T)f_1(T) + \cdots + g_s(T)f_s(T)]v \in V_1 \oplus \cdots \oplus V_s$, por lo que $V = V_1 \oplus \cdots \oplus V_s$ y T actúa como el escalar a_i en V_i. □

Una transformación lineal $T \in \mathrm{End}(V)$ es **nilpotente** cuando existe $r \in \mathbb{N}$ tal que $T^r = 0$.

Observaciones. Las siguientes afirmaciones son claras:

1. Si T es simultáneamente semisimple y nilpotente, entonces $T = 0$.

2. El polinomio característico $\det(\lambda - T)$ de T es un polinomio en λ con coeficientes que a su vez son polinomios en las coordenadas de cualquier matriz $A = (a_{ij})$ asociada a T.

3. El discriminante $g(a_{ij})$ del polinomio característico de T es un polinomio en a_{ij} tal que la inecuación $g(a_{ij}) \neq 0$ define un conjunto de matrices semisimples en $M_n(k)$.

Una **bandera** en V es una cadena de subespacios:

$$0 = W_0 \subset W_1 \subset \cdots \subset W_n = V, \text{ tal que } \dim W_i = i \ \forall i.$$

Dada una base ordenada $\{u_1, ..., u_n\}$ de V, le asociamos una bandera de manera natural así: $W_i = ku_1 + \cdots + ku_i$. Decimos que T **estabiliza** a la bandera $W_1 \subset \cdots \subset W_n$ cuando $T(W_i) \subseteq W_i$ para todo i. Esto sucede si y sólo si la matriz de T con respecto a $\{u_1, ..., u_n\}$ es triangular:

$$\begin{pmatrix} * & & \mathbf{0} \\ & \ddots & \\ \mathbf{*} & & * \end{pmatrix}.$$

Teorema 4.38 *Toda T es triangulable, es decir, existe una bandera de V estable ante T.*

Demostración: Procedemos por inducción en $n = \dim V$, siendo claro el resultado para $n = 1$.

Como k es algebraicamente cerrado, el polinomio característico $f(X) = \det(X - T)$ se factoriza así:

$$\prod_{i=1}^{r} (X - a_i)^{n_i} \text{ con todo } a_i \in k.$$

Esto implica que $f(a_i) = \det(a_i - T) = 0$ y que la transformación lineal $a_i - T$ es singular, por lo que existe $0 \neq v \in V$ tal que $(a_i - T)v = 0$, es decir, $T(v) = a_i v$. Hemos exhibido un vector característico de T.

Así, T estabiliza la línea $W_1 = \langle v \rangle$ y actúa en V/W_1. Concluimos por la hipótesis inductiva. □

Teorema 4.39 *Sean A un conjunto de transformaciones lineales que conmutan entre sí y B un subconjunto de A consistente de elementos semisimples. Entonces existe una base de V con respecto a la cual los elementos de A son triangulares y los de B son diagonales.*

Demostración: Procedemos por inducción en $n = \dim V$, siendo claro el resultado para $n = 1$.

Si todos los elementos de A son escalares, entonces no hay nada que hacer. Suponemos esto falso y consideramos los dos casos siguientes:

Caso 1: No todos los elementos de B son escalares. Sea $T \in B$ no escalar y sean $V_i = \{v \in V \mid T(v) = \lambda_i v\} \neq 0$ tales que $V = V_1 \oplus \cdots \oplus V_r$. Aquí tenemos que todo $S \in A$ estabiliza a todo V_i, pues

$$v \in V_i \Rightarrow T(Sv) = S(Tv) = S(\lambda_i v) = \lambda_i(Sv) \Rightarrow Sv \in V_i.$$

Por la hipótesis inductiva, existen bases de los V_i tales que con respecto a ellas, los elementos de A son triangulares y los de B (incluyendo a T) son diagonales. Entonces, la unión es una base de V como en el enunciado.

Caso 2: Todos los elementos de B son escalares. Por tanto, ya son diagonales y es suficiente triangular simultáneamente los elementos de A.

Sean $T \in A$ no escalar y $V_\lambda = \{v \in V \mid T(v) = \lambda v\} \neq 0$; este último espacio existe por el teorema anterior. Cada elemento de A estabiliza a V_λ y actúa en V/V_λ. Por la inducción, existe una base de V/V_λ con respecto a la cual A es triangular. Esta base se levanta a V para formar junto con una base de V_λ, una base de V como en el enunciado. □

Teorema 4.40 (Descomposición de Jordan-Chevalley) *Supongamos dada una transformación lineal T.*

a) Existen $S, N \in \mathrm{End}(V)$ únicos, tales que S es semisimple, N es nilpotente, $T = S + N$ y $SN = NS$.

b) Existen polinomios $p(X), q(X) \in k[X]$ sin término constante, tales que $S = p(T)$ y $N = q(T)$. En particular, S y N conmutan con toda transformación lineal de V que conmute con T.

c) Si $U \subseteq W \subseteq V$ son subespacios vectoriales tales que $T(W) \subseteq U$, entonces $S(W) \subseteq U$ y $N(W) \subseteq U$.

d) Si $T_1, T_2 \in \mathrm{End}(V)$ conmutan entre sí, entonces la parte semisimple (nilpotente) de $T_1 + T_2$ es la suma de las partes semisimples (nilpotentes) de T_1 y de T_2.

En las condiciones del teorema, S es la **parte semisimple** de T, mientras que N es la **parte nilpotente** de T.

Demostración: Supongamos que $\det(X - T) = \prod_{i=1}^{r}(X - a_i)^{m_i}$. Sean $V_i = \ker((T - a_i)^{m_i})$. Así, los subespacios V_i forman suma directa.

Escribimos $f_i(X) = \det(X - T)/(X - a_i)^{m_i} \in k[X]$ para $1 \leq i \leq r$. Es inmediato que se tiene la igualdad de "ideales" $k[X] = (f_1(X), ..., f_r(X))$, por lo que existen polinomios $g_1, ..., g_r$ tales que $g_1 f_1 + \cdots + g_r f_r = 1$. Dado $v \in V$, podemos escribir $v = [g_1(T)f_1(T) + \cdots + g_r(T)f_r(T)]v \in V_1 \oplus \cdots \oplus V_r$, por lo que $V = V_1 \oplus \cdots \oplus V_r$. Además, T estabiliza a cada V_i; y ahí satisface $(T - a_i)^{m_i} = 0$.

Por el Teorema Chino del Resíduo, existe $p(X) \in k[X]$ tal que

$$\begin{aligned} p(X) &\equiv a_i (\mathrm{mod}(X - a_i)^{m_i}),\ \forall i \\ p(X) &\equiv 0 (\mathrm{mod}\, X) \end{aligned}$$

Sean $q(X) = X - p(X)$, $S = p(T)$ y $N = q(T)$. Entonces S es semisimple, porque estabiliza a cada V_i y ahí actúa como multiplicación escalar por a_i. Como $N = T - S$ también estabiliza a cada V_i y actúa como $T - a_i$, se ve que N es nilpotente.

Hemos demostrado b) y la existencia de a). Veamos la unicidad de a): Las igualdades $T = S + N = S' + N', SN = NS, S'N' = N'S'$ con S, S' semisimples y N, N' nilpotentes implican que $S - S' = N' - N$ es nilpotente por ser simultaneamente triangulables N y N' por una parte y semisimple por ser simultaneamente diagonalizables S y S' por la otra. Por tanto, $S - S' = N' - N = 0$.

Las afirmaciones de c) son inmediatas.

Veamos d): Si $T_1 = S_1 + N_1$ y $T_2 = S_2 + N_2$ son descomposiciones de Jordan, entonces $T_1 + T_2 = (S_1 + S_2) + (N_1 + N_2)$ con $S_1 + S_2$ semisimple, que conmuta con $N_1 + N_2$ nilpotente, por lo que esta es una descomposición de Jordan. □

Algunos Grupos Lineales

Supongamos que V es un espacio vectorial de dimensión n sobre un campo arbitrario k; y que $T : V \to V$ es una transformación lineal. Presentaremos algunos importantes grupos multiplicativos de matrices en $M_n(k)$.

Sea $\mathbb{B} = \{A = (a_{ij}) \in M_n(k) \mid \det A \neq 0,\ a_{ij} = 0\ \forall\ i < j\}$. Este es el conjunto de las transformaciones lineales $T \in \operatorname{Aut} V$ que estabilizan la bandera $W_1 \subset \cdots \subset W_n$, obtenida de una base $\{v_1, ..., v_n\}$ de V así: $W_i = \langle v_1, ..., v_i \rangle$. La representación matricial de T es con respecto a esta base. Claramente, $\mathbb{B} < GL(V)$. Decimos que $\mathbb{B}$ es un **subgrupo de Borel** de $GL(V)$.

Sea $\mathbb{U} = \{A = (a_{ij}) \in B \mid a_{ii} = 1\ \forall i\}$. Cada $T \in \mathbb{B}$ actúa en cada cociente W_{i+1}/W_i de manera natural; $\mathbb{U}$ consiste de los elementos de $\mathbb{B}$ para los que todas estas acciones inducidas son la identidad.

Decimos que una transformación lineal es **unipotente** cuando todos sus valores característicos están en el campo k y son iguales a 1. Como los valores característicos de una matriz triangular son los elementos de la diagonal principal, vemos que $\mathbb{U}$ consiste de los elementos unipotentes de $\mathbb{B}$. Decimos que $\mathbb{U}$ es el **grupo unipotente** de $\mathbb{B}$.

La función $\varphi : \mathbb{B} \to (k^\star)^n$, dada por $\varphi(T) = (a_1, ..., a_n)$ cuando la acción de T en W_{i+1}/W_i es multiplicación escalar por a_{i+1}, es un morfismo suprayectivo de grupos, cuyo núcleo es $\mathbb{U}$. Por eso, $\mathbb{U} \lhd \mathbb{B}$.

Así, tenemos la sucesión exacta

$$1 \to \mathbb{U} \hookrightarrow \mathbb{B} \xrightarrow{\varphi} (k^\star)^n \to 1.$$

Existe otro morfismo $\psi : (k^\star)^n \to \mathbb{B}$, dado por

$$\psi(a_1, ..., a_n) = \operatorname{diag}(a_1, ..., a_n),$$

que satisface $\varphi \circ \psi = 1$, la identidad en $(k^\star)^n$. La imagen $\mathbb{T} = \psi[(k^\star)^n]$ es un **toro máximo** de $\mathbb{B}$.

Recordemos que E_{ij} tiene al número 1 en la posición ij y ceros en las otras posiciones; y que $x_{ij}(t) = 1 + tE_{ij}$, para $i \neq j$ y $t \in k$.

Definimos a los **grupos unipotentes de un parámetro** $\mathbb{U}_{ij}$ como $\operatorname{Im} x_{ij} = \{x_{ij}(t) \mid t \in k\}$, para cada $i \neq j$.

Teorema 4.41 *Los grupos $\mathbb{T}, \mathbb{U}, \mathbb{U}_{ij}$ y $\mathbb{B}$ tienen las siguientes propiedades:*
a) $\mathbb{B}$ *es un grupo soluble.*
b) $\mathbb{U}$ *es un grupo nilpotente.*
c) $\mathbb{U}_{ij} < \mathbb{U} \lhd \mathbb{B}$ *para $i > j$; y $\mathbb{T}$ normaliza a cada $\mathbb{U}_{ij}$.*
d) $\mathbb{B} = \mathbb{U} \rtimes \mathbb{T}$.
e) $\mathbb{U}$ *está generado por $\{\mathbb{U}_{ij} \mid i > j\}$.*
f) $\mathbb{U}_{ij} \cong k_+$, *el grupo aditivo de k.*
g) *Si $n \geq 2$ y $\circ(k) \geq 4$, entonces $\mathbb{U} = (\mathbb{B}, \mathbb{B})$, el grupo derivado de $\mathbb{B}$.*

Demostración: *f*) es consecuencia de la ecuación (4.10): $x_{ij}(t)\ x_{ij}(r) = x_{ij}(t + r), \forall\ t, r \in k$.

c) Claramente, $\mathbb{U}_{ij} < \mathbb{U}$, para $i > j$. Se tiene que $\mathbb{U}$ es normal en $\mathbb{B}$ porque $\mathbb{U} = \ker \varphi$. Se infiere que $\mathbb{T}$ normaliza a cada $\mathbb{U}_{ij}$ de la relación $\operatorname{diag}(a_1, ..., a_n) x_{ij}(t) [\operatorname{diag}(a_1, ..., a_n)]^{-1} = x_{ij}(a_i a_j^{-1} t)$.

d) Claramente, $\mathbb{T}$ y $\mathbb{U}$ generan a $\mathbb{B}$. Como $\mathbb{U} \lhd \mathbb{B}$ y $\mathbb{T} \cap \mathbb{U} = 1$, se tiene que $\mathbb{B}$ es el producto semidirecto indicado.

e) Procedemos por inducción en n, siendo claro el resultado para $n = 1$. Suponemos que $n \geq 2$ y que $A = (a_{ij}) \in \mathbb{U}$. Entonces

$$x_{n1}(-a_{n1}) \cdots x_{21}(-a_{21}) A = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \mathbf{0} \\ \vdots & * & \ddots & \\ 0 & & & 1 \end{pmatrix},$$

que nos permite concluir por la hipótesis inductiva.

g) Como $\mathbb{U} \lhd \mathbb{B}$ es tal que $\mathbb{B}/\mathbb{U} \cong \mathbb{T}$ es Abeliano, se tiene que $(\mathbb{B}, \mathbb{B}) \subseteq \mathbb{U}$. Recíprocamente, debido a *e*), es suficiente ver que $\mathbb{U}_{ij} \subseteq (\mathbb{B}, \mathbb{B})$ para todo $i > j$: Cuando $i > j + 1$, existe un entero ℓ tal que $i > \ell > j$, por lo que $x_{ij}(t) = (x_{i\ell}(t), x_{\ell j}(1))$ para todo $t \in k$.

Cuando $i = j + 1$, concentramos nuestra atención en las submatrices de las columnas y renglones $j, j + 1$. Esto nos reduce al caso $n = 2$, donde $\circ(k) \geq 4 \Rightarrow$ existe $c \in k^\star$ tal que $c^2 \neq 1$. El cálculo siguiente demuestra que $\mathbb{U}_{21} \subseteq (\mathbb{B}, \mathbb{B})$:

$$\begin{pmatrix} c^{-1} & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \begin{pmatrix} c & 0 \\ 0 & c^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ t(c^2 - 1) & 1 \end{pmatrix}.$$

b) Para cada $r \in \mathbb{N}$, definimos $\mathbb{U}_r = \langle \mathbb{U}_{ij} \mid i > j + r \rangle$, de manera que $\mathbb{U}_r < \mathbb{U}$ con $\mathbb{U}_r = 1$ para r lo suficientemente grande; por ejemplo, tan grande que no haya índices $i > j + r$. Así, tenemos la cadena de subgrupos

$$\mathbb{U} = \mathbb{U}_0 \supseteq \mathbb{U}_1 \supseteq \cdots \supseteq \mathbb{U}_r \supseteq \cdots \supseteq \mathbb{U}_n = 1$$

para la que las inclusiones $L_{i+1}U = (U, L_iU) \subseteq (U, U_i) \subseteq U_{i+1}$ se pueden obtener inductivamente, demostrando que $\mathbb{U}$ es un grupo nilpotente.

a) Como $\mathbb{U}$ es soluble y normal en $\mathbb{B}$ con $\mathbb{B}/\mathbb{U}$ Abeliano, se tiene que $\mathbb{B}$ es soluble. □

Ejercicios

1. Sea G un grupo multiplicativo finito de matrices en $M_n(\mathbb{C})$. Demuestre que toda matriz en G es diagonalizable.

2. Sea $D : V \to V$ la derivada, donde V es el espacio vectorial de polinomios de grado $\leq n$ sobre un campo, para $n > 1$. Demuestre que no existe ninguna base de V con respecto a la cual D sea diagonal.

3. Sea $A \in M_2(\mathbb{R})$ con $\det A < 0$. Demuestre que A es similar con una matriz diagonal.

4. Sea $T : V \to V$ una transformación lineal de un espacio vectorial de dimensión finita V sobre un campo k. Demuestre que el polinomio mínimo y el polinomio característico de T tienen los mismos factores irreducibles, aunque tal vez no con la misma multiplicidad.

5. Encuentre la descomposición de Jordan de

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ & & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix} \in M_n(\mathbb{C}).$$

6. Sea $T = S + N$ una descomposición de Jordan, con S semisimple y N nilpotente.

   a) Demuestre que T es invertible si y sólo si S lo es.

   b) En la situación de *a*), demuestre que S^{-1} es un polinomio en T.

   c) Demuestre que T es unipotente si y sólo si $T - 1$ es nilpotente.

   d) **(Descomposición de Jordan multiplicativa)** Suponga que T es invertible. Demuestre que existen S semisimple y U unipotente únicas, tales que $T = SU = US$, donde S y U son polinomios en T.

7. Con la notación del texto, demuestre que si k es infinito, entonces

   a) $Z(\mathbb{T}) = \mathbb{T}$.

   b) $N(\mathbb{T})/Z(\mathbb{T}) \cong S_n$, el grupo simétrico en n símbolos.

4.8 Conmutatividad de Matrices

Sea V un espacio vectorial de dimensión finita sobre un campo k. El anillo $\operatorname{End} V$ está muy lejos de ser conmutativo, por lo que tenemos el problema importante de encontrar qué transformaciones conmutan con una dada.

Iniciamos la discusión con un ejemplo sencillo. Consideremos la matriz

$$A = \left(\begin{array}{cc|ccc} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right)$$

asociada a la transformación lineal $T : V \to V$ del espacio vectorial V sobre el campo k, con respecto a la base $\{u_1, ..., u_5\}$. Observamos que A está en forma canónica racional, que si $R = k[X]$ actúa en V de manera que X actúe como T, entonces la estructura de R-módulo de V es:

$$V = Ru_1 \oplus Ru_3, \text{ con } \operatorname{an}(u_1) = X^2 + 1, \ \operatorname{an}(u_3) = X^3 + X,$$

donde $u_2 = Xu_1, u_4 = Xu_3, u_5 = X^2u_3$.

Una función lineal $L : V \to V$ que conmute con T quedará determinada por su acción en u_1 y en u_3, pues $L(u_2) = L(Xu_1) = L(Tu_1) = TL(u_1)$, así como $L(u_4) = L(Xu_3) = TL(u_3)$ y $L(u_5) = L(X^2u_3) = T^2L(u_3)$; pero $L(u_1)$ y $L(u_3)$ no pueden ser arbitrarios, tienen que satisfacer las condiciones $(X^2 + 1)L(u_1) = 0$ y $(X^3 + X)L(u_3) = 0$.

Escribiendo $L(u_1) = \alpha(X)u_1 + \beta(X)u_3$, vemos que $\alpha(X)$ puede ser un polinomio arbitrario, mientras que $\beta(X)$ debe satisfacer $X \mid \beta(X)$. De manera análoga, si $L(u_3) = \gamma(X)u_1 + \delta(X)u_3$, vemos que $\gamma(X)$ y $\delta(X)$ pueden ser polinomios arbitrarios, pues $(X^3 + X)(\alpha(X)u_1 + \beta(X)u_3) = 0$ en todo caso. Así, existen escalares $a_1, ..., a_9 \in k$ tales que

$$\begin{aligned} L(u_1) &= a_1u_1 + a_2Xu_1 + a_3Xu_3 + a_4X^2u_3 \\ L(u_3) &= a_5u_1 + a_6Xu_1 + a_7u_3 + a_8Xu_3 + a_9X^2u_3, \end{aligned}$$

por lo que, con respecto a esta base, L está representada por la matriz

$$\left(\begin{array}{cc|ccc} a_1 & a_2 & 0 & a_3 & a_4 \\ -a_2 & a_1 & 0 & -a_4 & a_3 \\ \hline a_5 & a_6 & a_7 & a_8 & a_9 \\ -a_6 & a_5 & 0 & a_7 - a_9 & a_8 \\ -a_5 & -a_6 & 0 & -a_8 & a_7 - a_9 \end{array}\right).$$

La matriz anterior fue construida a partir de los primeros renglones de cada bloque, para después completarse de la única forma posible. Las matrices que conmutan con A forman un espacio vectorial de dimensión 9.

Teorema 4.42 (Frobenius) *Sea $T : V \to V$ una transformación lineal con factores invariantes $d_1(X), ..., d_r(X)$ de grados $n_1 \leq \cdots \leq n_r$ y con $V = Rv_1 \oplus \cdots \oplus Rv_r$. Sea $Z(T)$ el conjunto de las transformaciones lineales de V que conmutan con T, entonces*

a) $Z(T)$ consiste de las transformaciones $L : V \to V$ tales que $L(v_i) = \alpha_{i1}(T)v_1 + \cdots + \alpha_{ir}(T)v_r$, para $1 \leq i \leq r$, donde $\alpha_{ij}(X)$ es un polinomio arbitrario si $j \leq i$; y que satisface $\alpha_{ij}(X) \in (d_j/d_i)R$, si $j > i$.

b) $\dim Z(T) = \sum_{i=1}^{r}(2r - 2i + 1)n_i = n_r + 3n_{r-1} + 5n_{r-2} + \cdots$

Demostración: *a)* Observemos que a T le corresponde una matriz en forma canónica racional:

$$A = \left(\begin{array}{c|c|c} * & & \\ \hline & \ddots & \\ \hline & & * \end{array}\right) \tag{4.12}$$

con bloques diagonales de tamaños $n_1 \leq \cdots \leq n_r$.

Tenemos que $V = k[X]v_1 \oplus \cdots \oplus k[X]v_r$, donde $\dim k[X]v_i = n_i \ \forall i$. La matriz A está asociada a la siguiente base de V:

$$\{v_1, Xv_1, ..., X^{n_1-1}v_1, ..., v_r, Xv_r, ..., X^{n_r-1}v_r\}. \tag{4.13}$$

Si B es una matriz que conmuta con A, entonces B está asociada a una transformación lineal $L : V \to V$, que está totalmente determinada por su acción en $\{v_1, v_2, ..., v_r\}$; pero que está sujeta a las condiciones $d_i(X)L(v_i) = 0$, para $1 \leq i \leq r$.

Fijemos un índice i y propongamos a $\alpha_{i1}(X)v_1 + \cdots + \alpha_{ir}(X)v_r \in V$ como candidato para ser $L(v_i)$. Dado que $d_1(X) \mid \cdots \mid d_r(X)$, la condición $d_i(X)L(v_i) = 0$ significa que $\alpha_{ij}(X)$ puede ser arbitrario si $j \leq i$, mientras que para $j > i$ se requiere que

$$(d_j/d_i) \mid \alpha_{ij}.$$

Así, para $j > i$, el polinomio α_{ij} puede ser cualquier elemento del k-espacio vectorial $(d_j/d_i)k[X]v_j$, cuya dimensión es n_i, mientras que para $j \leq i$, el polinomio α_{ij} puede ser cualquier elemento del k-espacio vectorial $k[X]v_j$, cuya dimensión es n_j. Así,

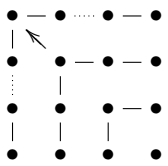
$$\mathcal{Z}_{ij} = \begin{cases} (d_j/d_i)k[X]v_j, \text{ si } j > i \\ k[X]v_j, \text{ si } j \leq i \end{cases} \quad \Rightarrow \quad \dim \mathcal{Z}_{ij} = \min\{n_i, n_j\}.$$

Situados en un bloque en la diagonal principal, al movernos en línea recta hacia la derecha o hacia abajo, se preserva la dimensión del espacio $\mathcal{Z}_{ij}$ asociado a cada bloque.

De lo anterior, obtenemos un procedimiento para generar todas las matrices que conmutan con A, así como para calcular la dimensión del espacio vectorial que forman tales matrices:

Fijamos la base (4.13) y tomamos la partición (4.12) para todas las matrices a considerar, donde el bloque ij corresponde al espacio $\mathcal{Z}_{ij}$. Llenamos el primer renglón de cada bloque cumpliendo con los requisitos que acabamos de precisar; y completamos cada bloque: En este paso, solamente aparecen en cada reglón, combinaciones lineales de las coordenadas del primer renglón del bloque en que nos encontremos, pues $LX^m v_i = T^m L v_i$, $\forall m$.

b) La dimensión de $Z(T)$ es una suma de términos de forma n_i, el número de estos términos es igual al número de bloques en (4.12), que es r^2. El término n_r aparece una vez, pues corresponde al bloque del extremo inferior derecho; el término n_{r-1} aparece tres veces, pues corresponde al gancho siguiente como en la figura; el término n_{r-2} aparece cinco veces, pues corresponde al gancho siguiente, etc.



La anterior manera de agrupar términos es la expresada en *b*). Notemos que $\sum_{i=1}^{r}(2r-2i+1) = 1+3+5+\cdots = r^2$; y que esta igualdad está ilustrada en la figura de arriba. □

Corolario 4.43 *Para una transformación lineal $T : V \to V$, las siguientes condiciones en son equivalentes:*

a) *T es cíclica.*

b) *$Z(T) = \{L \in \operatorname{End} V \mid LT = TL\}$ consiste de los polinomios en T.*

c) *$Z(T)$ es un conjunto conmutativo de transformaciones lineales.*

Demostración: *a*) $\Rightarrow$ *b*): Si T es cíclica, entonces T tiene un único factor invariante de grado $n = \dim V$. El Teorema de Frobenius dice que $\dim Z(T) = n$; pero el conjunto de los polinomios en T está contenido en $Z(T)$ y ya tiene dimensión n.

b) $\Rightarrow$ *c*): Esto es claro.

c) $\Rightarrow$ *a*): Supongamos T no cíclica, entonces $V = k[X]v_1 \oplus \cdots \oplus k[X]v_r$ con $r \geq 2$ y $\operatorname{an}(v_1) \supseteq \cdots \supseteq \operatorname{an}(v_r)$. Es suficiente exhibir $L_1, L_2 \in Z(T)$ tales que $L_1 L_2 \neq L_2 L_1$.

Para construir a L_1, L_2, podemos decidir que ambas actúen como la identidad en $k[X]v_3 \oplus \cdots \oplus k[X]v_r$ y así pasar al caso $r = 2$. Esto nos permite suponer que $V = k[X]u \oplus k[X]v$ con $\operatorname{an}(u) \supseteq \operatorname{an}(v)$.

Las siguientes condiciones determinan transformaciones lineales únicas $L_1, L_2 \in Z(T)$:

$L_1(u)$	$=$	0	$L_1(v)$	$=$	u
$L_2(u)$	$=$	u	$L_2(v)$	$=$	0

Aquí, $L_1L_2(v) = 0$; pero $L_2L_1(v) = L_2(u) = u$. □

Corolario 4.44 *El centro de $M_n(k)$, con k un campo arbitrario, consiste de las matrices escalares.*

Demostración: Claramente, toda matriz escalar es central. Recíprocamente, si $A = (a_{ij})$ es central, entonces A conmuta con

$$B = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

y con B^t. Como B y B^t son cíclicas, tenemos que A es un polinomio en B y también es un polinomio en B^t. Así, A es triangular superior e inferior, con $a_{11} = \cdots = a_{nn}$, por lo que es claro que A es escalar. □

Teorema 4.45 (del Doble Centralizador) *Si $T : V \to V$ es una transformación lineal, entonces $Z(Z(T))$ consiste de los polinomios en T.*

Demostración: Claramente, todo polinomio en T conmuta con todo elemento de $Z(T)$.

Recíprocamente, supongamos que $V = k[X]v_1 \oplus \cdots \oplus k[X]v_r$ cumple con $\mathrm{an}(v_1) \supseteq \cdots \supseteq \mathrm{an}(v_r)$, al exigir que X actúe como T.

Por el Teorema 4.42 a), sabemos que existen transformaciones lineales $P_{ij} \in Z(T)$ para toda pareja de índices $i \leq j$, tales que $P_{ij}(v_t) = \delta_{tj}v_i$.

Si $L \in Z(Z(T))$ y $L(v_i) = \sum_{i=1}^{r} \alpha_{ij}(T)v_j$, entonces los cálculos

$$L(v_i) = L(P_{ii}v_i) = P_{ii}(Lv_i) = P_{ii}(\sum_{i=1}^{r} \alpha_{ij}(T)v_j) = \alpha_{ii}(T)v_i,$$
$$L(v_i) = L(P_{ir}v_r) = P_{ir}(Lv_r) = P_{ir}(\sum_{i=1}^{r} \alpha_{rj}(T)v_j) = \alpha_{rr}(T)v_i$$

demuestran que $L = \alpha_{rr}(T)$. □

Ejercicios

1. Sea k algebraicamente cerrado. Calcule

$$\min\{\dim Z(A) \mid A \in M_n(k) \text{ semisimple}\}.$$

2. Encuentre todas las matrices en $M_5(\mathbb{Q})$ que conmutan con

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

3. Sea A nilpotente. Demuestre que $\dim Z(A)$ es mínima (entre las dimensiones de centralizadores de matrices nilpotentes) si y sólo si el polinomio mínimo de A es X^n.

4. Calcule el valor mínimo de $\{\dim Z(A) \mid A \in M_n(k)\}$, para A nilpotente con polinomio mínimo $\neq X^n$.

4.9 Formas Bilineales y Cuadráticas

Sea V un espacio vectorial de dimensión finita sobre un campo k. Una **forma simétrica bilineal** o **producto interno** de V es una función $B : V \times V \to k$ tal que

1. $B(au+bv, w) = aB(u, w)+bB(v, w)$, para todos $a, b \in k$; $u, v, w \in V$.

2. $B(u, v) = B(v, u)$, para todos $u, v \in V$.

Sea $\{u_1, ..., u_n\}$ una base de V. Esta elección le asigna una matriz a la forma B así:

$$A = (a_{ij}) \in M_n(k), \text{ donde } a_{ij} = B(u_i, u_j).$$

La matriz A es **simétrica**: $a_{ij} = a_{ji}$, para toda pareja de índices.

Escribamos (u, v) en lugar de $B(u, v)$ y consideremos la base $\{u_1^*, ..., u_n^*\}$ del espacio dual V^*, donde $u_i^*(u_j) = \delta_{ij}$ para $1 \leq i, j \leq n$.

Proposición 4.46 *La función $h : V \to V^*$ dada por $h(u)(v) = (u, v)$, para todos $u, v \in V$, es lineal; y le corresponde la matriz A con respecto a las bases duales elegidas.*

Demostración: Escribimos $h(u_i) = \sum_{j=1}^n c_{ij}u_j^*$ para cada i con $c_{ij} \in k$; y calculamos: $c_{ir} = (\sum_{j=1}^n c_{ij}u_j^*)(u_r) = h(u_i)(u_r) = (u_i, u_r) = a_{ir}$. □

El núcleo de h es el **radical** de la forma. Se dice que la forma es **no singular** cuando su radical es cero. Esto sucede si y sólo si h es biyectiva.

Ejemplo. Sean $k = \mathbb{R}$ y $A = 1$. En este caso, $\mathbb{R}^n$ provisto de la forma $(u, v) = a_1b_1 + \cdots + a_nb_n$, si $u = a_1u_1 + \cdots + a_nu_n$ y $v = b_1u_1 + \cdots + b_nu_n$ es el **espacio euclideano**.

En general, si $x = x_1u_1 + \cdots + x_nu_n,\ y = y_1u_1 + \cdots + y_nu_n \in V$ con $x_i, y_j \in k$ para todos i, j, se tiene que

$$(x, y) = \sum_{i,j=1}^{n} a_{ij}x_iy_j = (x_1, ..., x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Teorema 4.47 *Si a la forma bilineal B le corresponde la matriz $A = (a_{ij}) \in M_n(k)$ con respecto a la base $\{u_1, ..., u_n\}$, entonces con respecto a la base $\{v_1, ..., v_n\}$ dada por $v_i = \sum_{j=1}^{n} p_{ij}u_j$, para $1 \leq i \leq n$, le corresponde la matriz $C = PAP^t$, donde $P = (p_{ij})$ y P^t es la transpuesta de P.*

Demostración: Escribimos $C = (c_{ij})$ y calculamos

$$c_{ij} = (v_i, v_j) = (\sum_{r=1}^{n} p_{ir}u_r, \sum_{s=1}^{n} p_{js}u_s)$$
$$= \sum_{r,s=1}^{n} p_{ir}(u_r, u_s)p_{js} = \sum_{r,s=1}^{n} p_{ir}a_{rs}p_{js}. \quad \square$$

Decimos que u y v son **ortogonales** cuando $(u, v) = 0$. Dado $X \subseteq V$ un subconjunto arbitrario de V, definimos $X^\perp = \{v \in V \mid (x, v) = 0,\ \forall x \in X\}$. Es claro que $X^\perp$ siempre es un subespacio de V y que $V^\perp$ es el radical de la forma. Si U es un subespacio de V, decimos que $U^\perp$ es el **complemento ortogonal** de U.

Teorema 4.48 *Si U es un subespacio de V tal que $U \cap U^\perp = 0$, entonces $V = U \oplus U^\perp$.*

Demostración: Dado $v \in V$ arbitrario, es suficiente exhibir $u \in U$ y $w \in U^\perp$ tales que $v = u + w$.

Como $U \cap U^\perp$ es el radical de la forma restringida a U, vemos que la restricción es no singular. Por tanto, la función $\varphi : U \to U^*$ dada por $\varphi(x)(y) = (x, y)$ para $x, y \in U$ es suprayectiva.

El vector v dado produce un elemento $f \in U^*$ que actúa así: $f(x) = (v, x)$ para todo $x \in U$. Así que existe $u \in U$ tal que $(v, x) = (u, x),\ \forall x \in U$. Esto significa que $w = v - u \in U^\perp$. $\square$

Se dice que un vector v es **isotrópico** cuando $(v, v) = 0$.

Atención: A partir de este momento suponemos que la característica del campo k no es dos.

Teorema 4.49 *Siempre existe una base ortogonal de V, es decir, una base $\{u_1, ..., u_n\}$ tal que $(u_i, u_j) = 0$ si $i \neq j$.*

Demostración: Si la forma es idénticamente cero, no hay más que hacer. En caso contrario, existe un vector no isotrópico, pues la identidad

$$(u+v, u+v) = (u,u) + 2(u,v) + (v,v)$$

y la hipótesis caract $k \neq 2$ implican que si todo vector es isotrópico, entonces la forma es idénticamente cero.

Así, podemos suponer que u_1 no es isotrópico, para tener que $U = \langle u_1 \rangle$ satisface $U \cap U^\perp = 0$; y por el teorema anterior, $V = U \oplus U^\perp$.

Concluimos por inducción en $\dim V$, pues $\dim U^\perp < \dim V$; y $U^\perp$ tiene una base ortogonal, que junto con u_1 forma una base ortogonal de V. □

Este teorema admite la siguiente interpretación matricial:

Teorema 4.50 *Si $A \in M_n(k)$ es simétrica, entonces existe $P \in M_n(k)$ invertible tal que PAP^t es diagonal.*

Diremos que dos matrices $A, B \in M_n(k)$ son **congruentes** cuando exista $P \in M_n(k)$ invertible tal que $B = PAP^t$. Es inmediato que congruencia es una relación de equivalencia, que dos matrices congruentes tienen el mismo rango; y que si nos restringimos a matrices simétricas, dos matrices son congruentes si y sólo si están asociadas a una misma forma simétrica bilineal, donde la matriz P que las relaciona, expresa el cambio de base correspondiente.

Una **forma cuadrática** $Q : V \to k$ es una función polinomial, homogénea de segundo grado en las coordenadas de $v \in V$. Toda forma bilineal B da origen a una forma cuadrática Q así: $Q(v) = \frac{1}{2}B(v,v),\ \forall v \in V$.

Si V tiene un producto interno, diremos que V es una **suma directa ortogonal**, escrito $V = V_1 \perp \cdots \perp V_r$ cuando se tengan

$$V = V_1 \oplus \cdots \oplus V_r \ \text{ y } \ (v_i, v_j) = 0 \Leftarrow v_i \in V_i, v_j \in V_j, i \neq j.$$

Si $k = \mathbb{R}$, se dice que una forma bilineal B o su forma cuadrática asociada Q es **positiva definida** cuando $v \neq 0 \Rightarrow Q(v) > 0$. Se dice que B ó Q es **negativa definida** cuando $v \neq 0 \Rightarrow Q(v) < 0$.

Teorema 4.51 (de la Inercia de Sylvester) *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{R}$, provisto de una forma simétrica bilineal. Entonces existe una descomposición de V como suma directa ortogonal*

$$V = V_1 \perp V_2 \perp V_3,$$

tal que la forma restringida a V_1 es positiva definida, en V_2 es negativa definida y en V_3 es cero. La descomposición puede no ser única; pero las dimensiones de los sumandos ortogonales sí lo son.

Demostración: Partimos de una base ortogonal $\{u_1, ..., u_n\}$, que existe por el Teorema 4.49.

Escribiendo $(u_i, u_i) = a_i$ y con el propósito de crear una nueva base ortogonal $\{v_1, ..., v_n\}$ tal que $(v_i, v_i) \in \{1, -1, 0\}$ para todo i, definimos los vectores v_i como sigue:

$$\begin{aligned} \frac{u_i}{\sqrt{a_i}}, &\text{ si } a_i > 0, \\ \frac{u_i}{\sqrt{-a_i}}, &\text{ si } a_i < 0, \\ u_i, &\text{ si } a_i = 0. \end{aligned}$$

Sean $V_1 = \langle v_i \mid (v_i, v_i) = 1\rangle, V_2 = \langle v_i \mid (v_i, v_i) = -1\rangle, V_3 = \langle v_i \mid (v_i, v_i) = 0\rangle$. Resulta que $V = V_1 \perp V_2 \perp V_3$ es como en el enunciado. Si $V = W_1 \perp W_2 \perp W_3$ es otra descomposición como en el enunciado, entonces

$$\dim V_3 = \dim W_3 = \dim V - \text{rango } A,$$

para cualquier matriz A asociada al producto interno.

Sea $f : V \to V_1$ la proyección lineal. La restricción de f a W_1 tiene núcleo $U = W_1 \cap (V_2 \perp V_3)$, donde $0 \neq v \in U \Rightarrow (v, v) \leq 0$ y también $(v, v) > 0$. Así, $U = 0$, por lo que W_1 se inyecta en V_1. Esto demuestra que $\dim W_1 \leq \dim V_1$, la igualdad $\dim W_1 = \dim V_1$ se obtiene por simetría. □

La interpretación matricial del teorema anterior es la siguiente:

Teorema 4.52 *Si $A \in M_n(\mathbb{R})$ es simétrica, entonces existe $P \in M_n(\mathbb{R})$ invertible tal que $PAP^t = \text{diag}(1, ..., 1, -1, ..., -1, 0, ..., 0)$.*

Dos matrices diagonales con elementos en $\{1, -1, 0\}$ son congruentes si y sólo si tienen el mismo número de "unos", "menos unos" y "ceros" en la diagonal.

El teorema anterior produce formas canónicas ante congruencia para matrices reales simétricas.

Dada una matriz $A \in M_n(k)$, en muchos casos podemos encontrar su forma canónica ante congruencia sin tener que encontrar una matriz P que efectúe la congruencia. El **discriminante de un producto interno** es el determinante de cualquier matriz asociada. Para un producto interno no singular, este es un invariante como elemento de $k^\star/(k^\star)^2$.

Los **menores líderes** de una matriz $A \in M_n(k)$ son:

$$\delta_r = \Delta^{1,2,...,r}_{1,2,...,r}(A), \text{ para } 1 \leq r \leq n.$$

Si la matriz A está asociada a un producto interno con respecto a la base $\{u_1, ..., u_n\}$, entonces sus menores líderes son los discriminantes del producto interno restringido a los subespacios $\langle u_1, ..., u_i\rangle$, por lo que A es congruente con $\text{diag}(\delta_1, \delta_2/\delta_1, ..., \delta_n/\delta_{n-1})$.

Ejemplo. Consideremos la matriz simétrica $A \in M_3(\mathbb{R})$ siguiente:

$$A = \begin{pmatrix} 4 & 2 & 7 \\ 2 & -11 & 6 \\ 7 & 6 & 1 \end{pmatrix}, \text{ donde } \delta_1 = 4,\ \delta_2 = -48 \text{ y } \delta_3 = 515.$$

Así, existe $P \in M_3(\mathbb{R})$ invertible con $PAP^t = \text{diag}(1, -1, -1)$.

El siguiente resultado es inmediato.

Teorema 4.53 *Si $A \in M_n(\mathbb{R})$ es simétrica con menores líderes distintos de cero, entonces*

a) A es positiva definida si y sólo si $\delta_r > 0$ para todo r.

b) A es negativa definida si y sólo si $\delta_{2r+1} < 0$ y $\delta_{2r} > 0$ para todo r.

Sea V un espacio euclideano, es decir, $V = \mathbb{R}^n$, provisto de un producto interno positivo definido. La **longitud** de un vector v es $|v| = \sqrt{(v, v)}$. La **distancia** entre dos vectores u y v es $d(u, v) = |u - v|$.

Proceso de ortogonalización de Gram-Schmidt

A partir de un subconjunto linealmente independiente $\{u_1, ..., u_n\}$ del espacio euclideano V, es posible obtener una base ortogonal $\{v_1, ..., v_n\}$ del espacio $\langle u_1, ..., u_n\rangle$ tal que la bandera asociada a ambos conjuntos sea la misma:

$$\langle u_1, ..., u_i\rangle = \langle v_1, ..., v_i\rangle, \ \forall i.$$

El primer paso es muy fácil: Tomamos $v_1 = u_1$.

Resolvemos para c_{12} la siguiente ecuación:

$$0 = (v_1, u_2 - c_{12}v_1) = (v_1, u_2) - c_{12}(v_1, v_1),$$

lo cual es posible porque $(v_1, v_1) > 0$. Escribimos $v_2 = u_2 - c_{12}v_1$ para tener $(v_1, v_2) = 0$ con $\{v_1, v_2\}$ linealmente independiente.

En el siguiente paso resolvemos para c_{13} y c_{23} las siguientes ecuaciones:

$$0 = (v_1, u_3 - c_{13}v_1) = (v_1, u_3) - c_{13}(v_1, v_1),$$
$$0 = (v_2, u_3 - c_{23}v_2) = (v_2, u_3) - c_{23}(v_2, v_2),$$

lo cual es posible porque $(v_1, v_1) > 0$ y $(v_2, v_2) > 0$.

Escribimos $v_3 = u_3 - c_{13}v_1 - c_{23}v_2$ para tener $(v_1, v_3) = (v_2, v_3) = 0$ con $\{v_1, v_2, v_3\}$ linealmente independiente. Continuando de esta manera se llega a una base ortogonal como se deseaba, pues para cada i, tenemos que

$$u_i - v_i \in \langle u_1, ..., u_{i-1}\rangle.$$

Volúmenes m-dimensionales

Dados m vectores $u_1, ..., u_m$ con $m \leq n$ en un espacio euclideano $\mathbb{R}^n$, el **m-paralelepípedo** que generan es

$$P(u_1, ..., u_m) = \{v = \sum_{i=1}^{m} c_i u_i \in \mathbb{R}^n \mid 0 \leq c_i \leq 1\}.$$

Sea $\mathcal{P}_{m,n}$ la colección de todos los m-paralelepípedos en $\mathbb{R}^n$ así generados.

Queremos definir una función volumen $\mathbf{v}_m : \mathcal{P}_{m,n} \to \mathbb{R}$. Si el conjunto $\{u_1, ..., u_m\}$ es ortogonal, queremos tener

$$\mathbf{v}_m(P(u_1, ..., u_m)) = \prod_{i=1}^{m} |u_i|.$$

Si el conjunto $\{u_1, ..., u_m\}$ es linealmente dependiente, queremos tener

$$\mathbf{v}_m(P(u_1, ..., u_m)) = 0.$$

Dado un conjunto linealmente independiente $\{u_1, ..., u_m\}$, lo ortogonalizamos à la Gram-Schmidt hasta obtener $\{v_1, ..., v_m\}$. Interpretamos cada expresión $u_i = (u_i - v_i) + v_i$ como la descomposición ortogonal de u_i en su componente $u_i - v_i$ en la "base" $\langle u_1, ..., u_{i-1}\rangle$ de $P(u_1, ..., u_i)$, más la "altura" v_i. Basados en esta interpretación, definimos

$$\mathbf{v}_m(P(u_1, ..., u_m)) = \prod_{i=1}^{m} |v_i|.$$

Teorema 4.54 *a) Si* $u_i = \sum_{j=1}^{n} a_{ij}\epsilon_j$ *para* $1 \leq i \leq m$, *donde* $\{\epsilon_1, ..., \epsilon_n\}$ *es la base natural de* $\mathbb{R}^n$ *y* $A = (a_{ij}) \in M_{m\times n}(\mathbb{R})$, *entonces*

$$\mathbf{v}_m(P(u_1, ..., u_m)) = \sqrt{\det AA^t}.$$

b) Si $T : V \to V$ *es lineal, entonces*

$$\mathbf{v}_m(P(Tu_1, ..., Tu_m)) = |\det T|\mathbf{v}_m(P(u_1, ..., u_m)).$$

Demostración: Procedemos por inducción en m. Cuando $m = 1$, tenemos $\mathbf{v}_1(P(u_1)) = |v_1| = |u_1| = \sqrt{(u_1, u_1)} = \sqrt{u_1 u_1^t} = \sqrt{AA^t} = \sqrt{\det AA^t}$.

Sea A_1 la submatriz de A obtenida al eliminar el último renglón. Podemos suponer que

$$\mathbf{v}_{m-1}(P(u_1, ..., u_{m-1})) = \sqrt{\det A_1 A_1^t}.$$

Como $v_m - u_m \in \langle u_1, ..., u_{m-1}\rangle$, existe una matriz C, que es producto de matrices elementales de primer tipo, tal que

$$CA = \begin{pmatrix} A_1 \\ v_m \end{pmatrix} \quad \text{y} \quad \det C = \det C^t = 1.$$

Esto nos permite realizar los siguientes cálculos:

$$CAA^tC^t = \begin{pmatrix} A_1 \\ v_m \end{pmatrix} \begin{pmatrix} A_1^t & v_m^t \end{pmatrix}$$

$$= \begin{pmatrix} A_1A_1^t & A_1v_m^t \\ v_mA_1^t & v_mv_m^t \end{pmatrix} = \begin{pmatrix} A_1A_1^t & 0 \\ 0 & v_mv_m^t \end{pmatrix};$$

$$\det AA^t = (\det A_1 A_1^t)|v_m|^2 = \prod_{i=1}^{m} |v_i|^2,$$

de donde se obtiene a). La afirmación b) es inmediata. □

Ejercicios

1. Sea V un espacio vectorial de dimensión finita sobre $\mathbb{R}$, provisto de un producto interno sin vectores isotrópicos distintos de cero. Demuestre que el producto interno es positivo definido o negativo definido.

2. Sea k algebraicamente cerrado. Demuestre que dos matrices A, B en $M_n(k)$, simétricas, son congruentes si y sólo si tienen el mismo rango. Encuentre formas canónicas para este caso.

3. Sean V un espacio vectorial de dimensión finita provisto de un producto interno y $\mathcal{B} = \{w_1, ..., w_r\} \subset V$ tal que $\det A \neq 0$, donde $A = (a_{ij}) \in M_r(k)$ es tal que $a_{ij} = (w_i, w_j)$. Demuestre que $\mathcal{B}$ es linealmente independiente.

4. a) Demuestre que el área de un triángulo cuyos vértices son $(0,0,0)$, (a_1, a_2, a_3) y (b_1, b_2, b_3) es

$$\frac{1}{2}\sqrt{\det AA^t}, \quad \text{donde} \quad A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

   b) Demuestre que el volumen de un tetraedro con vértices en el origen y en los puntos $u_1, u_2, u_3 \in \mathbb{R}^n$ es

$$\frac{1}{6}\sqrt{\det AA^t}, \quad \text{donde} \quad A = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

5. Sea V un espacio vectorial con un producto interno. Sea u un vector no isotrópico. Definimos la **reflección** $T_u : V \to V$ así:

$$T_u(v) = -v + 2\frac{(v,u)}{(u,u)}u.$$

   La función $-T_u$ es la **simetría determinada por** u. Demuestre que
   a) $T_u(u) = u$ y que $T_u(w) = -w$ para $w \in \langle u \rangle^{\perp}$.
   b) $T_u^2 = 1$.
   c) T_u preserva el valor de los productos internos.

6. Demuestre las siguientes propiedades de la longitud y de la distancia:
   a) $|cv| = |c| \cdot |v|$ si $c \in \mathbb{R}$.
   b) $|(u,v)| \leq |u| \cdot |v|$.
   c) $|u+v| \leq |u| + |v|$.
   d) $d(u,w) \leq d(u,v) + d(v,w)$.

4.10 Formas Alternas

En esta sección suponemos que k es un campo de característica distinta de dos. Sea V un espacio vectorial de dimensión finita sobre k. Una **forma alterna** en V es una función bilineal

$$f : V \times V \to k, \text{ tal que } f(v, v) = 0, \text{ para todo } v \in V.$$

Escribimos (u, v) en lugar de $f(u, v)$ y observamos que $0 = (u+v, u+v) = (u, u) + (u, v) + (v, u) + (v, v)$ implica la propiedad llamada **antisimetría**:

$$(u, v) = -(v, u), \text{ para todos } u, v \in V.$$

Si a la forma alterna le corresponde la matriz $A = (a_{ij}) \in M_n(k)$ con respecto a la base $\{u_1, ..., u_n\}$, donde $a_{ij} = (u_i, u_j)$, entonces A es una **matriz alterna**, es decir $a_{ij} = -a_{ji}$ para todos i, j; y además, $a_{ii} = 0$ para todo i. Esto es, $A^t = -A$.

Observaciones. Las siguientes afirmaciones son inmediatas.

1. La matriz A corresponde, con respecto a bases duales elegidas, a la función lineal $h : V \to V^*$ dada por $h(u)(v) = (u, v)$, para todos $u, v \in V$.

2. La forma f es no degenerada si y sólo si h es biyectiva, equivalentemente, si y sólo si $\det A \neq 0$.

3. Si la forma f es no degenerada, entonces n es par, pues $\det A = \det A^t = (-1)^n \det A$.

4. El rango de toda forma o matriz alterna es par.

Teorema 4.55 *Sea V un espacio vectorial de dimensión finita provisto de una forma alterna no degenerada. Entonces existe una base $\{v_1, v_2, ..., v_{2n}\}$ tal que $(v_{2i-1}, v_{2i}) = 1 = -(v_{2i}, v_{2i-1})$ para $1 \leq i \leq n$; y tal que $(v_r, v_s) = 0$ si $\{r, s\} \neq \{2i-1, 2i\}$.*

Demostración: Hacemos inducción en n, siendo claro el caso $n = 0$.

Elegimos $v_1 \in V$ de manera arbitraria. Como existen vectores no ortogonales a v_1, elegimos $v_2 \in V$ con $(v_1, v_2) = 1$.

Sea $W = \langle v_1, v_2 \rangle$. Afirmamos que $W \cap W^\perp = 0$. Esto es porque

$$av_1 + bv_2 \in W^\perp \Rightarrow (v_1, av_1 + bv_2) = b = 0 \text{ y } (av_1 + bv_2, v_2) = a = 0.$$

Como $W \cap W^\perp = 0$, tenemos que $V = W \oplus W^\perp$ con $\dim W^\perp < \dim V$, por el Teorema 4.48. Concluimos por la inducción. □

Corolario 4.56 *Todas las formas alternas no degeneradas en un espacio vectorial V son equivalentes ante* $\operatorname{Aut} V$.

La interpretación matricial de los resultados anteriores es la siguiente:

Teorema 4.57 *Si $A \in M_n(k)$ es una matriz alterna invertible, entonces existe otra matriz invertible $P \in M_n(k)$ tal que*

$$PAP^t = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}. \tag{4.14}$$

Llamémosle A_0 a la matriz en la derecha de la ecuación (4.15). Esta es una matriz alterna asociada a una base como en el Teorema 4.55, le llamaremos la **matriz alterna canónica**. Es claro que $\det A_0 = 1$.

Dada una matriz alterna A, tenemos B invertible tal que $A = BA_0B^t$, por lo que $\det A = (\det B)(\det B^t) = (\det B)^2$. Definimos el **Pfaffiano** de A, escrito $\operatorname{Pf} A$, como $\det B$, de manera que $\det A = (\operatorname{Pf} A)^2$.

Observaciones.

1. La definición anterior parece ambigua, porque no tenemos unicidad de B. La propiedad $\det A = (\operatorname{Pf} A)^2$ reduce la ambigüedad al factor ± 1. El próximo teorema demostrará que la ambigüedad no existe.

2. De nuestra definición, es claro que $\operatorname{Pf}(CAC^t) = (\det C)(\operatorname{Pf} A)$.

Ejemplos. En los casos $n = 2$ y $n = 4$, tenemos que

$$\det \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} = a^2,$$

$$\det \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{pmatrix} = (a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23})^2.$$

Consideremos el anillo de polinomios $R = \mathbb{Z}[a_{ij}]$ en las $n(n-1)/2$ variables a_{ij} con $1 \leq i < j \leq n$. Escribiendo $a_{ii} = 0$ para $1 \leq i \leq n$; y también $a_{ij} = -a_{ji}$ para $i > j$, definimos la **matriz alterna genérica** $A = (a_{ij}) \in M_n(R)$.

Existe $B \in M_n(\mathbb{Q}(a_{ij}))$ tal que $A = BA_0B^t$, entonces $\det B = \operatorname{Pf} A$ es el **Pfaffiano genérico**, que satisface $\det A = (\operatorname{Pf} A)^2 \in \mathbb{Z}[a_{ij}]$. La unicidad de la factorización en $\mathbb{Z}[a_{ij}]$ implica que $\operatorname{Pf} A \in \mathbb{Z}[a_{ij}]$.

El grupo **octaédrico** Oct_n consiste de los elementos $w \in S_{2n}$ que al actuar en $\{\pm 1, \pm 2, ..., \pm n\}$, cumplen con $w(-i) = -w(i)$, para $1 \leq i \leq n$. Claramente, $\circ(Oct_n) = n!2^n$. También hay una acción natural de Oct_n en $\{\{1,2\},\{3,4\},...,\{2n-1,2n\}\}$.

Teorema 4.58 *Si $A \in M_{2n}(k)$ es una matriz alterna invertible, entonces*

$$\begin{aligned}\operatorname{Pf} A &= \sum_{\sigma \in S_{2n}/Oct_n} (-1)^\sigma \prod_{i=1}^{n} a_{\sigma(2i-1),\sigma(2i)} \\ &= \frac{1}{n!2^n} \sum_{\sigma \in S_{2n}} (-1)^\sigma \prod_{i=1}^{n} a_{\sigma(2i-1),\sigma(2i)}.\end{aligned}$$

Demostración: Sea V el espacio vectorial sobre k de dimensión $2n$ subyacente. Dada $T \in \operatorname{Aut} V$, consideramos su acción inducida en el álgebra exterior $\bigwedge V$.

Concentramos nuestra atención en $\bigwedge^2 V$, donde se encuentran dos elementos que corresponden a las matrices A_0 y A, o a sus formas bilineales asociadas:

$$f_0 = \sum_{i=1}^{n} (v_{2i-1} \wedge v_{2i}), \quad f = \sum_{i<j} a_{ij}(v_i \wedge v_j).$$

Supongamos que $(\bigwedge^2 T)(f_0) = f$, esto equivale a requerir la igualdad $A = PA_0P^t$, si P es la matriz asociada a T. Entonces

$$(\bigwedge^{2n} T)\left(\frac{f_0^n}{n!}\right) = \frac{f^n}{n!};$$

pero efectuando los cálculos obtenemos

$$\frac{1}{n!} f_0^n = \frac{1}{n!} (\sum_{i=1}^{n} v_{2i-1} \wedge v_{2i})^n = v_1 \wedge v_2 \wedge \cdots \wedge v_{2n},$$

de manera que

$$(\bigwedge^{2n} T)\left(\frac{f_0^n}{n!}\right) = (\det T) v_1 \wedge v_2 \wedge \cdots \wedge v_{2n}.$$

Por otra parte,

$$\frac{f^n}{n!} = \frac{1}{n!} (\sum_{i<j} a_{ij} v_i \wedge v_j)^n =$$

$$\frac{1}{n!2^n} \sum_{\sigma \in S_{2n}} (-1)^\sigma \prod_{i=1}^{n} a_{\sigma(2i-1),\sigma(2i)} (v_1 \wedge v_2 \wedge \cdots \wedge v_{2n}),$$

de donde llegamos a la conclusión. □

Ejemplo. Sea A_{ij} la submatriz $(n-2) \times (n-2)$ obtenida a partir de A omitiendo los renglones y columnas i, j. Sea $\alpha_{ij} = (-1)^{i+j-1} \operatorname{Pf} A_{ij}$.

Aprovechando el razonamiento anterior, descomponemos $f = f_1 + f_2$, escribiendo $f_1 = \sum_{j=2}^{n} a_{1j}(v_1 \wedge v_j)$ y $f_2 = \sum_{1<i<j} a_{ij}(v_i \wedge v_j)$. Calcularemos $\operatorname{Pf} A = \det T$ como sigue:

$$\begin{aligned}(\operatorname{Pf} A)v_1 \wedge \cdots \wedge v_{2n} &= \frac{1}{n!}f^n = \frac{1}{n!}(f_1 + f_2)^n \\ &= \frac{1}{n!}\binom{n}{1}(f_1 \wedge f_2^{n-1}) = f_1 \wedge \frac{1}{(n-1)!}f_2^{n-1} \\ = [\sum_{j=2}^{n} a_{1j}(v_1 \wedge v_j)] \wedge [\sum_{j=2}^{n} &(\operatorname{Pf} A_{1j})\widehat{v}_1 \wedge v_2 \wedge \cdots \wedge \widehat{v}_j \wedge \cdots \wedge v_{2n}] \\ &= [\sum_{j=2}^{n}(-1)^j a_{1j}(\operatorname{Pf} A_{1j})]v_1 \wedge \cdots \wedge v_{2n},\end{aligned}$$

En el cálculo anterior, usamos el Teorema del Binomio, pues f_1 y f_2 conmutan entre sí, y con cualquier otra cosa, por ser de grado par. Obtuvimos

$$\operatorname{Pf} A = \sum_{j=2}^{n} a_{1j}\alpha_{1j}.$$

El grupo multiplicativo $\{T \in \operatorname{Aut} V \mid (\bigwedge^2 T)(f_0) = f_0\}$ es el **grupo simpléctico** Sp_{2n}.

Corolario 4.59 *Si $T \in Sp_{2n}$, entonces* $\det T = 1$.

Demostración: Si $T \in Sp_{2n}$, entonces $(\bigwedge^2 T)(f_0) = f_0$. De esta manera, $(\bigwedge^{2n} T)$ estabiliza a $(1/n!)f_0^n = v_1 \wedge v_2 \wedge \cdots \wedge v_{2n}$; pero la acción de $(\bigwedge^{2n} T)$ en $v_1 \wedge v_2 \wedge \cdots \wedge v_{2n}$ es multiplicación por $\det T$. Así, $\det T = 1$. □

Ejercicios

1. Demuestre que si se intercambian los reglones r y s de una matriz alterna A; y simultáneamente se intercambian las columnas r y s, entonces $\operatorname{Pf} A$ cambia de signo.

2. Demuestre que si se multiplican simultáneamente el renglón r y la columna r de una matriz alterna A por un escalar c, entonces $\operatorname{Pf} A$ queda multiplicado por c.

3. Demuestre que $Sp_2 = SL_2$.

4. Dada una matriz $A = (a_{ij}) \in M_n(k)$ con n par y renglones $A_1, ..., A_n$; definimos un producto para los renglones $A_i \cdot A_j = c_{ij} \in k$ así:

$$c_{ij} = \begin{vmatrix} a_{i1} & a_{i2} \\ a_{j1} & a_{j2} \end{vmatrix} + \begin{vmatrix} a_{i3} & a_{i4} \\ a_{j3} & a_{j4} \end{vmatrix} + \cdots + \begin{vmatrix} a_{i,n-1} & a_{i,n} \\ a_{j,n-1} & a_{j,n} \end{vmatrix}$$

   Demuestre que la matriz $C = (c_{ij})$ es alterna y que $\operatorname{Pf} C = \det A$.

5. Demuestre que el Pfaffiano genérico es irreducible en $\mathbb{Z}[a_{ij}]$.

4.11 Formas Hermitianas

Sea V un espacio vectorial de dimensión finita sobre $\mathbb{C}$. Para $\alpha = a+bi \in \mathbb{C}$ con $a, b \in \mathbb{R}$, escribimos la conjugación compleja así: $\overline{\alpha} = a-bi$. Una **forma Hermitiana** o **producto interno sesquilineal** en V es una función

$$S : V \times V \to \mathbb{C} \quad \text{tal que}$$

1. $S(au+bv, w) = aS(u, w)+bS(v, w)$, para todos $a, b \in \mathbb{C}$; $u, v, w \in V$.
2. $S(u, v) = \overline{S(v, u)}$, para todos $u, v \in V$.

Cuando S queda entendida, escribimos (u, v) en lugar de $S(u, v)$. Al elegir una base $\{u_1, ..., u_n\}$ de V, a la forma S se le asocia una matriz así:

$$A = (a_{ij}) \in M_n(\mathbb{C}), \text{ donde } a_{ij} = (u_i, u_j).$$

La matriz A es **Hermitiana**, es decir $\overline{a_{ij}} = a_{ji}$, para toda pareja de índices. En particular, los elementos de la diagonal son reales.

Durante el resto de la sección supondremos que V tiene una forma Hermitiana fija positiva definida, es decir, que $0 \neq v \in V \Rightarrow (v, v) \in \mathbb{R}, (v, v) > 0$. En estas condiciones diremos que V es un **espacio unitario**. Aquí definimos la **longitud** de un vector v como $|v| = \sqrt{(v, v)}$ y la **distancia** entre dos vectores u y v como $d(u, v) = |u - v|$.

Observaciones. Las siguientes afirmaciones o son inmediatas o admiten demostraciones similares a las ya dadas para afirmaciones semejantes, usamos la notación de arriba.

1. Si $x = x_1u_1 + \cdots + x_nu_n$, $y = y_1u_1 + \cdots + y_nu_n$, entonces

$$(x, y) = \sum_{i,j=1}^{n} a_{ij}x_i\overline{y_j} = (x_1, ..., x_n)A \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}.$$

2. La longitud y la distancia cumplen las propiedades del Ejercicio 4.9.6.

3. Con respecto a una nueva base $\{v_1, ..., v_n\}$ de V, relacionada con la base anterior así: $v_i = \sum_{j=1}^{n} p_{ij}u_j$ $\forall i$, a S le corresponde la matriz PAP^*, donde $P^* = (q_{ij})$ con $q_{ij} = \overline{p_{ji}}$ es la **adjunta Hermitiana** de $P = (p_{ij})$.

4. Existe una base **ortonormal** de V, esto es, una base $\{w_1, ..., w_n\}$, tal que $(w_i, w_j) = \delta_{ij}$.

5. Existe $P \in M_n(\mathbb{C})$ invertible tal que PAP^* es la identidad.

6. La función $f : V \to V^*$ dada por $f(u)(v) = (v, u)$, para $u, v \in V$, es semilineal y biyectiva. Decir que f es **semilineal** significa que

$$f(u_1 + u_2) = f(u_1) + f(u_2),\ f(\alpha u) = \overline{\alpha} f(u),\ \forall u, u_1, u_2 \in V,\ \alpha \in \mathbb{C}.$$

Dados $T : V \to V$ y $v \in V$, podemos definir $h \in V^*$ así: $h(u) = (T(u), v)$, para todo $u \in V$. La Observación 6 garantiza la existencia de un único vector $w \in V$ tal que $h(u) = (u, w)$. Esto nos permite definir una función $T^* : V \to V$, dada por $T^*(v) = w$ en la situación anterior. La función T^* es la **adjunta Hermitiana** de T. Notemos que se cumple la igualdad

$$(T(u), v) = (u, T^*(v)),\ \forall\ u, v \in V,\ T \in \operatorname{End} V. \tag{4.15}$$

Proposición 4.60 *La adjunta Hermitiana T^* de una transformación T tiene las siguientes propiedades:*

a) T^* es lineal.

b) Si la matriz de T es $A = (a_{ij})$ con respecto a una base ortonormal de V y T^* tiene asociada a la matriz $B = (b_{ij})$ con respecto a la misma base, entonces $a_{ij} = \overline{b_{ji}}$.

c) $T^{**} = T$.

d) $(S + T)^* = S^* + T^*$, para toda S lineal.

e) $(\alpha T)^* = \overline{\alpha} T^*$, para toda $\alpha \in \mathbb{C}$.

f) $(ST)^* = T^* S^*$, para toda S lineal.

Demostración: *a)* Partiendo de (4.16), tenemos que

$$(u, T^*(av + bw)) = (Tu, av + bw) = \overline{a}(Tu, v) + \overline{b}(Tu, w) =$$
$$\overline{a}(u, T^* v) + \overline{b}(u, T^* w) = (u, aT^* v + bT^* w),\ \forall a, b \in \mathbb{C};\ u, v, w \in V.$$

Como V es un espacio unitario, se obtiene la linealidad de T^*.

b) Sea $\{u_1, ..., u_n\}$ una base ortonormal de V con $T(u_i) = \sum_{j=1}^n a_{ij} u_j$. Estamos suponiendo que $T^*(u_r) = \sum_{s=1}^n b_{rs} u_s$, para $1 \leq r \leq n$. Entonces:

$$a_{ij} = (Tu_i, u_j) = (u_i, T^* u_j) = (u_i, \sum_{s=1}^n b_{js} u_s) = \overline{b_{ji}}.$$

Las afirmaciones *c), d)* y *e)* son consecuencias inmediatas de *b)*. Para ver *f)*, calculamos: $(u, (ST)^* v) = (STu, v) = (Tu, S^* v) = (u, T^* S^* v)$. □

Dada $A = (a_{ij}) \in M_n(\mathbb{C})$, definimos su **adjunta Hermitiana** como la matriz $B = (b_{ij})$ tal que $b_{ij} = \overline{a_{ji}}$, para todos $1 \leq i, j \leq n$. Escribimos la adjunta Hermitiana de A así: A^*.

Teorema 4.61 *Las siguientes condiciones en $T : V \to V$ son equivalentes:*

a) $(Tv, Tv) = (v, v)$ para todo $v \in V$.

b) $(Tu, Tv) = (u, v)$ para todos $u, v \in V$.

c) T envía una base ortonormal a otra.

d) $T^* = T^{-1}$.

e) $TT^* = 1 = T^* T$.

Demostración: La equivalencia $d) \Leftrightarrow e)$ es clara, por lo que veremos las implicaciones $a) \Rightarrow b) \Rightarrow c) \Rightarrow e) \Rightarrow a)$.

$a) \Rightarrow b)$: Como $(T(u+v), T(u+v)) = (u+v, u+v)$, tenemos que

$$(Tu, Tv) + (Tv, Tu) = (u, v) + (v, u). \tag{4.16}$$

También tenemos $(Tu, T(-iv)) + (T(-iv), Tu) = (u, -iv) + (-iv, u)$, es decir, $i(Tu, Tv) - i(Tv, Tu) = i(u, v) - i(v, u)$, lo cual nos da

$$(Tu, Tv) - (Tv, Tu) = (u, v) - (v, u) \tag{4.17}$$

De (4.17) y (4.18) obtenemos $(Tu, Tv) = (u, v)$ para todos $u, v \in V$.

$b) \Rightarrow c)$: Si $\{u_1, ..., u_n\}$ es una base ortonormal de V y $T(u_i) = v_i$ para toda i, entonces $(v_i, v_j) = (Tu_i, Tu_j) = (u_i, u_j) = \delta_{ij}$.

$c) \Rightarrow e)$: Dada una base ortonormal $\{u_1, ..., u_n\}$, tenemos que

$$\delta_{ij} = (u_i, u_j) = (Tu_i, Tu_j) = (u_i, T^*Tu_j) \Rightarrow u_j = T^*Tu_j, \ \forall j.$$

Esto implica que $TT^* = 1 = T^*T$.

$e) \Rightarrow a)$: Si $v \in V$, entonces $(v, v) = (v, 1v) = (v, T^*Tv) = (Tv, Tv)$. □

Una transformación lineal T es **unitaria** cuando satisface las condiciones del teorema anterior. Se dice que T es **Hermitiana** cuando $T = T^*$, que es **antihermitiana** cuando $T = -T^*$, o que es **normal** cuando $TT^* = T^*T$. Se definen estas mismas nociones para matrices complejas cuadradas, de manera análoga. Así, resulta que la matriz asociada con respecto a una base ortonormal, a una transformación lineal unitaria (resp. Hermitiana, antihermitiana o normal) es unitaria (resp. Hermitiana, antihermitiana o normal).

Observaciones. Si $A = (a_{ij}) \in M_n(\mathbb{R})$, entonces

1. A es Hermitiana si y sólo si A es simétrica.
2. A es antihermitiana si y sólo si A es antisimétrica.
3. A es normal si y sólo si A conmuta con su transpuesta.
4. A es unitaria si y sólo si $AA^t = 1 = A^tA$, es decir, $A^t = A^{-1}$.

Una matriz $A \in M_n(\mathbb{R})$ tal que $A^t = A^{-1}$ se llama **ortogonal**.

Teorema 4.62 *Sea V un espacio unitario.*

a) Si $T \in \operatorname{End} V$ es Hermitiana y W es un subespacio de V estable ante T, entonces $T(W^\perp) \subseteq W^\perp$.

b) Si $\mathcal{A} \subseteq \operatorname{End} V$ satisface $T \in \mathcal{A} \Rightarrow T^* \in \mathcal{A}$ y si $T(W) \subseteq W$ para un subespacio W de V, entonces $T(W^\perp) \subseteq W^\perp$, $\forall\ T \in \mathcal{A}$.

c) Si $T \in \operatorname{End} V$ es Hermitiana, entonces V admite una base ortonormal formada por vectores característicos de T.

d) Si $\mathcal{A}$ es un conjunto de transformaciones lineales Hermitianas de V que conmutan entre sí, entonces V admite una base ortonormal formada por vectores característicos simultáneos de todas las $T \in \mathcal{A}$.

Demostración: *a*) es consecuencia inmediata de *b*).

b): Si $u \in W,\ v \in W^{\perp},\ T \in \mathcal{A}$, entonces $(Tv, u) = (v, T^*u) = 0$, porque $T^* \in \mathcal{A} \Rightarrow T^*u \in W$.

c) es consecuencia inmediata de *d*).

d): Por el Teorema 4.39, existe un vector característico v común a toda $T \in \mathcal{A}$. Si $W = \langle v \rangle$, entonces $W^{\perp}$ es estable ante toda $T \in \mathcal{A}$, gracias a *b*). Definimos $v_1 = v/\sqrt{(v,v)}$ y obtenemos por inducción en $\dim V$ una base ortonormal $\{v_2, ..., v_n\}$ de $W^{\perp}$, con todo v_i vector característico de toda $T \in \mathcal{A}$, de manera que $\{v_1, ..., v_n\}$ es una base ortonormal de V. □

Como la transición de una base ortonormal a otra, se da por medio de una matriz unitaria, tenemos la siguiente versión matricial del resultado anterior:

Corolario 4.63 *Dado un conjunto conmutativo de matrices Hermitianas $\mathcal{C}$, existe una matriz unitaria U tal que $UAU^* = UAU^{-1}$ es diagonal para toda $A \in \mathcal{C}$.*

Teorema 4.64 *Sea V un espacio unitario.*

a) Si $T : V \to V$ es normal, entonces V admite una base ortonormal de vectores característicos de T.

b) Dada una matriz normal N, existe una matriz unitaria U tal que $UNU^* = UNU^{-1}$ es diagonal.

c) Para T normal, T^* es un polinomio en T con coeficientes en $\mathbb{C}$.

d) Si $\mathcal{A}$ es un conjunto conmutativo de transformaciones normales, entonces V admite una base ortonormal formada por vectores característicos simultáneos de todas las $T \in \mathcal{A}$.

e) Dado un conjunto conmutativo de matrices normales $\mathcal{C}$, existe una matriz unitaria U tal que $UAU^* = UAU^{-1}$ es diagonal para toda $A \in \mathcal{C}$.

Demostración: *a*) Sean $T_1 = T + T^*$ y $T_2 = iT - iT^*$. Por ser T normal, tenemos que $T_1 = T_1^*, T_2 = T_2^*$ y que $T_1T_2 = T_2T_1$. Por tanto, V admite una base ortonormal de vectores característicos de T_1 y T_2; pero tenemos que $T = (1/2)(T_1 - iT_2)$. Así, estos vectores también son vectores característicos de T.

b) es la versión matricial de *a*).

c) Si a T se le asocia la matriz $\text{diag}(a_1, ..., a_n)$ con respecto a una base ortonormal de V, tenemos que a T^* se le asocia la matriz $\text{diag}(\overline{a_1}, ..., \overline{a_n})$ con respecto a la misma base.

El Teorema de Interpolación de Lagrange garantiza que existe un polinomio $p(X) \in \mathbb{C}[X]$ de grado menor o igual que $n-1$, tal que $p(a_i) = \overline{a_i}$, para $1 \leq i \leq n$. Entonces $p(T) = T^*$.

d) Sea $\mathcal{C} = \mathcal{A} \cup \{T^* \mid T \in \mathcal{A}\}$. Entonces *c*) garantiza que $\mathcal{C}$ es un conjunto conmutativo, por lo que existe un vector característico v común a todo $T \in \mathcal{A}$. El Teorema 4.62 *b*) implica que $\langle v\rangle^\perp$ es estable ante todo $T \in \mathcal{C}$. Definimos $v_1 = v/\sqrt{(v,v)}$ y obtenemos por inducción en $\dim V$ una base ortonormal $\{v_2, ..., v_n\}$ de $\langle v\rangle^\perp$, con todo v_i vector característico de toda $T \in \mathcal{C}$, de manera que $\{v_1, ..., v_n\}$ es una base ortonormal de V como la buscada.

e) es la versión matricial de *d*). □

Corolario 4.65 *a*) *Los valores característicos de toda transformación lineal T Hermitiana son reales.*

b) *Los valores característicos λ de toda transformación lineal unitaria T satisfacen $\lambda\overline{\lambda} = 1$.*

c) *Sea T una transformación normal. Si sus valores característicos son reales, entonces T es Hermitiana. Si sus valores característicos λ satisfacen $\lambda\overline{\lambda} = 1$, entonces T es unitaria.*

Demostración: Por ser T normal, en cualquier caso, le asociamos una matriz diagonal con respecto a una base ortonormal de V. Entonces todo es claro. □

Cuando el campo es $\mathbb{R}$, podemos refinar nuestros resultados como sigue.

Teorema 4.66 (del Eje Principal)

a) *Si $A \in M_n(\mathbb{R})$ es simétrica, entonces existe P ortogonal tal que $PAP^t = PAP^{-1}$ es diagonal.*

b) *El espacio vectorial $V = \mathbb{R}^n$ admite una base ortonormal consistente de vectores característicos de una transformación lineal simétrica dada.*

Demostración: *a*) y *b*) son afirmaciones equivalentes que demostraremos simultáneamente.

Tenemos una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ del espacio euclideano de vectores renglón, definida así: $T(v) = vA$, donde A es la matriz simétrica dada. La acción de T puede extenderse a $\mathbb{C}^n$ con la misma definición.

Como T es Hermitiana, sus valores característicos distintos $\lambda_1, ..., \lambda_r$ son reales. Sea $p(X) = \prod_{i=1}^{r}(X - \lambda_i)^{n_i}$ el polinomio característico de T. Entonces tenemos la descomposición $V = V_1 \oplus \cdots \oplus V_r$, donde $V_i = \ker(T - \lambda_i)^{n_i}$, para $1 \leq i \leq r$.

A la transformación normal T le corresponde una matriz diagonal sobre $\mathbb{C}$, con respecto a una base ortonormal de $\mathbb{C}^n$. Por ello,

$$(T - \lambda_i)^{n_i} v = 0 \Rightarrow (T - \lambda_i)v = 0,$$

es decir, todo $v_i \in V_i$ es un vector característico de T.

Si $v_i \in V_i$ y $v_j \in V_j$ con $i \neq j$, entonces $\lambda_i \neq \lambda_j$; y por lo tanto

$$\lambda_j(v_i, v_j) = (v_i, Tv_j) = (Tv_i, v_j) = \lambda_i(v_i, v_j) \Rightarrow (v_i, v_j) = 0.$$

Así tenemos que $V = V_1 \perp \cdots \perp V_r$ es una suma directa ortogonal de espacios en los que T actúa como escalar real, cada uno de los cuales admite una base ortonormal, cuya unión es una base de V como en el enunciado. □

Teorema 4.67 *Sea $B \in M_n(\mathbb{R})$ una matriz ortogonal. Entonces existe P ortogonal con $PBP^t = PBP^{-1}$ de forma*

$$\begin{pmatrix} 1_r & & & & & & \\ & -1_s & & & & & \\ & & \cos\alpha_1 & \operatorname{sen}\alpha_1 & & & \\ & & -\operatorname{sen}\alpha_1 & \cos\alpha_1 & & & \\ & & & & \ddots & & \\ & & & & & \cos\alpha_t & \operatorname{sen}\alpha_t \\ & & & & & -\operatorname{sen}\alpha_t & \cos\alpha_t \end{pmatrix}$$

Demostración: Sea $A = B + B^t = B + B^{-1}$. Entonces A es simétrica y conmuta con B. Por el teorema anterior, existe una descomposición $V = V_1 \perp \cdots \perp V_r$, donde $v_i \in V_i \Rightarrow v_i A = \lambda_i v_i$, con $\lambda_1, ..., \lambda_r$ elementos distintos de $\mathbb{R}$.

La función lineal L definida como multiplicación derecha por la matriz B, estabiliza a cada V_i, donde actúa biyectivamente.

Como $v_i(B+B^{-1}) = \lambda_i v_i$ para $v_i \in V_i$, tenemos que $v_i(B^2 - \lambda_i B + 1) = 0$, lo que implica que el subespacio $W = \langle v_i, v_i B\rangle$ de V_i es invariante ante L; pero entonces el complemento ortogonal $W^\perp$ de W en V_i también es invariante ante L: Si $u \in W$ y $v \in W^\perp$ entonces $(vB, u) = (v, B^{-1}u) = 0$, porque $B^t = B^{-1}$ y porque L actúa biyectivamente en W.

Así, cada V_i es una suma directa ortogonal de subespacios W invariantes ante L de dimensión ≤ 2.

Sea T la transformación lineal dada por multiplicación derecha por B restringida al espacio W.

Si $\dim W = 1$, entonces T es ortogonal y actúa como escalar. Este debe ser ± 1.

Si $\dim W = 2$, entonces con respecto a cualquier base ortogonal de W, se le asocia a T una matriz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

con $a^2 + b^2 = 1 = c^2 + d^2$ y $ac + bd = 0$. Además, $T^2 - \lambda_i T + 1 = 0 \Rightarrow \det T = 1$, esto es $ad - bc = 1$. Estas ecuaciones admiten una solución de la forma $a = d = \cos\alpha,\ b = -c = \operatorname{sen}\alpha$. □

Notemos que toda transformación lineal ortogonal, o bien toda matriz ortogonal tiene determinante ± 1. Una **rotación** es una transformación ortogonal con determinante uno.

Sea $T : V \to V$ una transformación lineal Hermitiana de un espacio unitario V. Sabemos que los valores característicos de T son reales. Decimos que T es **positiva definida** (resp. **negativa definida**) cuando todo valor característico λ de T es positivo (resp. negativo).

Proposición 4.68 *Sea T una función lineal de un espacio unitario V.*
a) Si T es Hermitiana, entonces T es positiva definida si y sólo si $(Tv, v) > 0$ para todo $0 \neq v \in V$.
b) Si T es invertible, entonces TT^* es positiva definida.

Demostración: *a*) Existe un base ortonormal $\{v_1, ..., v_n\}$ de V formada por vectores característicos de T: Aquí, $T(v_i) = \lambda_i v_i$, para $1 \leq i \leq n$. Si $\lambda_i > 0$, para todo i, entonces cualquier vector $0 \neq v \in V$ puede escribirse como $v = \sum_{i=1}^{n} a_i v_i$, de manera que

$$(Tv, v) = (T\sum_{i=1}^{n} a_i v_i, \sum_{i=1}^{n} a_i v_i) = (\sum_{i=1}^{n} \lambda_i a_i v_i, \sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} \lambda_i a_i \overline{a_i} > 0.$$

Recíprocamente, si v es un vector característico de T con $T(v) = \lambda v$, tenemos que

$$\lambda(v, v) = (Tv, v) > 0 \Rightarrow \lambda > 0.$$

b) Claramente TT^* es Hermitiana. Dado $0 \neq v \in V$, se tiene que $(TT^*v, v) = (T^*v, T^*v) > 0$, porque T invertible $\Rightarrow T^*$ invertible. □

Proposición 4.69 *Toda transformación lineal positiva definida T admite una única raíz cuadrada positiva definida, que es un polinomio en T.*

Demostración: Una vez que a T se le asocia una matriz $\text{diag}(\lambda_1, ..., \lambda_n)$ con respecto a una base ortonormal de V, con todo $\lambda_i > 0$, tomamos $\mu_i > 0$ con $\mu_i^2 = \lambda_i$ para cada i.

La transformación lineal H, asociada a la matriz $\text{diag}(\mu_1, ..., \mu_n)$ respecto a la base ortonormal anterior, satisface $H^2 = T$. La Interpolación de Lagrange produce un polinomio f tal que $f(\lambda_i) = \mu_i$, para todo i. Así, $f(T) = H$ conmuta con toda transformación lineal que conmute con T.

Si H_1 es otra raíz cuadrada positiva definida de T, entonces, al diagonalizar H_1, vemos que H_1 conmuta con T y con H, por lo que es posible diagonalizar simultáneamente H_1 y H. Así se ve que $H_1 = H$. □

Teorema 4.70 (Descomposición Polar) *Sea $T : V \to V$ una transformación lineal invertible de un espacio unitario V. Entonces existen H positiva definida y U unitaria, únicas; tales que $T = HU$.*

Demostración: Veamos la unicidad: $T = HU \Rightarrow T^* = U^*H$; y entonces $TT^* = HUU^*H = H^2$. Así, H es la única raíz cuadrada positiva definida de TT^*.

Veamos la existencia: Partiendo de H como acabamos de ver, definimos $U = H^{-1}T$, para tener $U^* = T^*(H^{-1})^* = T^*H^{-1}$, por lo que $UU^* = H^{-1}TT^*H^{-1} = 1$. Así, U es unitaria y $T = HU$. □

Ejercicios

1. Demuestre que para toda $A \in M_n(\mathbb{C})$, existe U unitaria tal que UAU^{-1} es triangular.

2. Demuestre que si para $A \in M_n(\mathbb{C})$, existe U unitaria tal que UAU^{-1} es diagonal, entonces A es normal.

3. Demuestre que si $A \in M_n(\mathbb{R})$ es invertible, entonces existen matrices B simétrica positiva definida y C ortogonal únicas, tales que $A = BC$.

4.12 Ejercicios Generales

1. Sea $A \in M_n(\mathbb{Q})$ con $\operatorname{tr} A = 0$. Demuestre que A es similar a una matriz con ceros en la diagonal principal.

2. Sea $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$. Demuestre que la derivada total de esta función, evaluada en $(a_1, ..., a_n)$; y escrita $(D\det)(a_1, ..., a_n)$ satisface

$$(D\det)(a_1, ..., a_n)(b_1, ..., b_n) = \sum_{i=1}^{n} \det(a_1, ..., a_{i-1}, b_i, a_{i+1}, ..., a_n).$$

3. Sean k un anillo conmutativo con uno y V un k-módulo libre de rango r. Dado $L \in \operatorname{End}_k V$, se genera una serie de transformaciones inducidas por L en distintos objetos construidos a partir de V.

   a) Demuestre que existe un único morfismo, llamado $\prod L$ que hace conmutativo a todo diagrama donde p_i es la proyección en el i-ésimo factor directo:

$$\begin{array}{ccc} V^n & \xrightarrow{p_i} & V \\ {\scriptstyle \prod L}\downarrow & & \downarrow{\scriptstyle L} \\ V^n & \xrightarrow{p_i} & V \end{array}$$

   b) Demuestre que los distintos $\prod L$ inducen morfismos únicos T^nL del álgebra tensorial T^nV, que hacen conmutativos a los diagramas

$$\begin{array}{ccc} V^n & \longrightarrow & T^nV \\ {\scriptstyle \prod L}\downarrow & & \downarrow{\scriptstyle T^nL} \\ V^n & \longrightarrow & T^nV \end{array}$$

y que inducen un único morfismo $T(L) : T(V) \to T(V)$, donde

$$T(V) = \coprod_{n\geq 0} T^n V.$$

c) Demuestre que los distintos $\prod L$ inducen morfismos únicos $S^n L$ del álgebra simétrico $S^n V$, que hacen conmutativos a los diagramas

$$\begin{array}{ccc} V^n & \longrightarrow & S^n V \\ {\scriptstyle \prod L}\downarrow & & \downarrow{\scriptstyle S^n L} \\ V^n & \longrightarrow & S^n V \end{array}$$

y que inducen un único morfismo $S(L) : S(V) \to S(V)$, donde

$$S(V) = \coprod_{n\geq 0} S^n V.$$

d) Demuestre que los distintos $\prod L$ inducen morfismos únicos $\bigwedge^n L$ del álgebra alternante $\bigwedge^n V$, que hacen conmutativos a los diagramas

$$\begin{array}{ccc} V^n & \longrightarrow & \bigwedge^n V \\ {\scriptstyle \prod L}\downarrow & & \downarrow{\scriptstyle \bigwedge^n L} \\ V^n & \longrightarrow & \bigwedge^n V \end{array}$$

y que inducen un único morfismo $\bigwedge L : \bigwedge V \to \bigwedge V$, donde

$$\bigwedge V = \coprod_{n=0}^{r} \bigwedge^n V.$$

e) Sean $L_1, L_2 \in \operatorname{End} V$. Demuestre que $T(L_1 \circ L_2) = T(L_1) \circ T(L_2)$ y que $T(1) = 1$. Enuncie y demuestre propiedades análogas para los otros morfismos inducidos.

f) Demuestre que $\bigwedge^n L$ es multiplicación por $\det L$.

g) Demuestre que a cada elemento de $\bigwedge^2 V$ se le asocia una matrix alterna A de manera natural. Al elegir una base de V, a L se le asocia la matriz C. Demuestre que entonces $\bigwedge^2 L$ corresponde a la transformación $A \mapsto CAC^t$.

4. Demuestre que dadas matrices $A \in M_{m\times n}$ y $B \in M_{n\times r}$, todo menor $t \times t$ de AB, con $t \leq m, n, r$; es una suma de productos $D_1 D_2$, donde D_1 es un menor $t \times t$ de A, mientras que D_2 lo es de B. Escriba la expresión precisa de este resultado.

Capítulo 5
Temas Complementarios

5.1 Teorema de la Base Normal

Dada una extensión finita de Galois F/k con $\mathrm{Gal}(F/k) = \{\sigma_1, ..., \sigma_n\}$, se dice que una base de F/k es una **base normal** cuando es de la forma $\{\sigma_1(w), ..., \sigma_n(w)\}$ para algún $w \in F$. Aquí veremos que toda extensión finita de Galois posee una base normal.

Proposición 5.1 *Sea F/k una extensión separable de campos de grado n y sean $\sigma_1, ..., \sigma_n$ los distintos k-morfismos $F \to \overline{k}$. Entonces un subconjunto $\mathcal{B} = \{a_1, ..., a_n\} \subseteq F$ es una base de F sobre k si y sólo si $\det(\sigma_i a_j) \neq 0$.*

Demostración: Teniendo el número correcto de elementos, $\mathcal{B}$ es una base de F/k si y sólo si $\mathcal{B}$ es linealmente independiente sobre k.

La ecuación lineal $a_1x_1 + \cdots + a_nx_n = 0$ tiene las mismas soluciones en k que el sistema de ecuaciones (que la incluye):

$$\begin{aligned} \sigma_1(a_1)x_1 + \cdots + \sigma_1(a_n)x_n &= 0 \\ &\vdots \\ \sigma_n(a_1)x_1 + \cdots + \sigma_n(a_n)x_n &= 0 \end{aligned} \tag{5.1}$$

pues las soluciones $a_1c_1 + \cdots + a_nc_n = 0$ con todo $c_i \in k$ generan automáticamente soluciones del sistema (5.1), al aplicar a esta expresión los distintos morfismos σ_i. Nuestra conclusión se obtiene al observar que la matriz de coeficientes del sistema de ecuaciones es $(\sigma_i a_j) \in M_n(F)$. □

Teorema 5.2 (de la Base Normal) *Sea F/k una extensión finita de Galois con $G = \mathrm{Gal}(F/k) = \{\sigma_1, ..., \sigma_n\}$. Entonces existe $w \in F$ tal que $\{\sigma_1(w), ..., \sigma_n(w)\}$ es una base de F/k.*

Demostración: Caso 1: k es finito. Aquí, el Teorema 3.50 *a*) afirma que G es cíclico, por lo que podemos suponer que $G = \langle\sigma\rangle$, con $\sigma^i = \sigma_i$ para $1 \leq i \leq n$.

El Teorema 3.59 afirma que el conjunto G es linealmente independiente sobre F y con mayor razón sobre k. Por tanto $\sigma \in \mathrm{End}_k F$ tiene polinomio

mínimo de grado $\geq n$; y la acción de σ hace de F un módulo cíclico. Esto significa precisamente que existe $w \in F$ tal que $\{\sigma_1(w), ..., \sigma_n(w)\}$ es una base de F/k.

Caso 2: k es infinito. Por el Corolario 3.37, existe un elemento primitivo $u \in F$ tal que $F = k(u)$. Sea $f(X) = \text{Polmin}(u, k)$. Escribimos $u_i = \sigma_i(u)$, para $1 \leq i \leq n$, observamos que $f(X) = \prod_{i=1}^{n}(X - u_i)$; y consideramos los siguientes polinomios:

$$q(X) = \frac{f(X)}{(X-u)f'(u)} \quad \text{y} \quad q_i(X) = \sigma_i[q(X)] = \frac{f(X)}{(X-u_i)f'(u_i)}$$

Los polinomios $q_i(X)$ están en $F[X]$ y tienen la siguiente propiedad:

$$q_i(u_j) = \delta_{ij}. \tag{5.2}$$

Ante la acción de G en $\{q_i(X) \mid 1 \leq \ i \leq n\}$, cada elemento $1 \neq \sigma \in G$ actúa sin puntos fijos.

Es inmediato que

$$i \neq j \ \Rightarrow \ q_i(X)q_j(X) \equiv 0 \pmod{f(X)}. \tag{5.3}$$

En la ecuación

$$\sum_{i=1}^{n} q_i(X) - 1 = 0, \tag{5.4}$$

el polinomio de la izquierda tiene grado $\leq n - 1$; pero (5.2) implica que tiene las n raíces $u_1, ..., u_n$; por lo que (5.4) es una identidad polinomial.

Multiplicamos (5.4) por $q_i(X)$ y usamos (5.3) para obtener

$$q_i(X)^2 \equiv q_i(X) \pmod{f(X)}. \tag{5.5}$$

Consideramos ahora la matriz $A = (\sigma_i\sigma_j[q(X)]) \in M_n(F[X])$, para la que obtenemos con ayuda de (5.3-5) que

$$AA^t \equiv 1 \pmod{f(X)}.$$

Por lo tanto, el polinomio $g(X) = \det A \in F[X]$ satisface

$$g(X)^2 \equiv 1 \pmod{f(X)}.$$

Así, $g(X)$ no es cero; y existe $b \in F$ tal que $g(b) \neq 0$. Escribiendo $w = q(b)$, esto significa que $\det(\sigma_i\sigma_j(w)) \neq 0$ y que $\{\sigma_1(w), ..., \sigma_n(w)\}$ es linealmente independiente sobre k. □

Proposición 5.3 *Sea F/k una extensión de campos con k infinito. Dado un polinomio $0 \neq p(X_1, ..., X_n) \in F[X_1, ..., X_n]$, existen $a_1, ..., a_n \in k$ tales que $p(a_1, ..., a_n) \neq 0$.*

Demostración: Procedemos por inducción en el número de variables n, siendo claro el caso $n = 1$, pues todo polinomio en una variable tiene un número finito de raíces. Escribimos nuestro polinomio

$$p(X_1, ..., X_n) = \sum_{i=0}^{r} q_i(X_1, ..., X_{n-1}) X_n^i,$$

donde se tiene que $q_i(X_1, ..., X_{n-1}) \in F[X_1, ..., X_{n-1}]$, para $0 \leq i \leq r$; y además $q_r(X_1, ..., X_{n-1}) \neq 0$.

Por la hipótesis inductiva, existen elementos $a_1, ..., a_{n-1} \in k$ tales que $q_r(a_1, ..., a_{n-1}) \neq 0$, por lo que

$$0 \neq \sum_{i=0}^{r} q_i(a_1, ..., a_{n-1}) X_n^i \in F[X_n].$$

De manera que existe $a_n \in k$ tal que $p(a_1, ..., a_n) \neq 0$. □

Teorema 5.4 (Independencia Algebraica de Morfismos) *Sea F/k una extensión separable de campos de grado n, con k infinito; y sea $\mathcal{A} = \{\sigma_1, ..., \sigma_n\}$ el conjunto de los distintos k-morfismos $F \to \overline{k}$. Entonces $\mathcal{A}$ es algebraicamente independiente sobre F.*

Demostración: Supongamos que $f(X_1, ..., X_n) \in F[X_1, ..., X_n]$ es tal que $f(\sigma_1(u), ..., \sigma_n(u)) = 0$, para todo $u \in F$.

Sea $\{u_1, ..., u_n\}$ una base de F sobre k. Efectuamos la transformación lineal T en las variables dada por $T(X_i) = \sum_{j=1}^{n} (\sigma_i u_j) X_j$, para $1 \leq i \leq n$, de manera que T es invertible al asociarse a la matriz $(\sigma_i u_j) \in M_n(F)$.

El polinomio $g(X_1, ..., X_n) = f(T(X_1), ..., T(X_n))$ satisface

$$\begin{aligned} g(a_1, ..., a_n) &= f(\Sigma_j (\sigma_1 u_j) a_j, ..., \Sigma_j (\sigma_n u_j) a_j) \\ &= f(\sigma_1(\Sigma_j a_j u_j), ..., \sigma_n(\Sigma_j a_j u_j)) = 0; \end{aligned}$$

siempre que $a_1, ..., a_n \in k$.

La proposición anterior implica que $g(X_1, ..., X_n) = 0$; pero entonces $f(X_1, ..., X_n) = 0$, ya que si L es la transformación lineal inversa de T, se tiene que $f(X_1, ..., X_n) = g(L(X_1), ..., L(X_n)) = 0$. □

Concluimos esta sección con una segunda demostración para el Caso 2 del Teorema de la Base Normal.

Demostración: Sea $R = k[X_1, ..., X_n]$. Consideramos la correspondencia $X_i \leftrightarrows \sigma_i$ entre variables y elementos de G. Sea $X \in M_n(R)$ la matriz que tiene a la variable $X_{i(j)}$ correspondiente a $\sigma_i \sigma_j$ en la posición ij. Escribimos $f(X_1, ..., X_n) = \det X$.

Como cada renglón y cada columna de la matriz X es una permutación del conjunto de las variables $X_1, ..., X_n$; vemos que $f(1, 0, ..., 0) = \pm 1$. Esto implica que $f(X_1, ..., X_n) \neq 0$.

Sabemos que G es algebraicamente independiente sobre F, por lo que $\det(\sigma_i \sigma_j) \neq 0$. Siendo F infinito, la Proposición 5.3 implica que existe

$w \in F$ tal que $\det((\sigma_i\sigma_j)(w)) \neq 0$; pero entonces $\{\sigma_1(w), ..., \sigma_n(w)\}$ es linealmente independiente sobre k. □

Ejercicios

1. Sea F/k una extensión finita de campos, de Galois con grupo de Galois $G = \mathrm{Gal}(F/k)$. Demuestre que F es un $k[G]$-módulo libre de rango uno.

2. Sea F/k una extensión finita de Galois con $G = \mathrm{Gal}(F/k)$ y con k infinito. Demuestre que existe $u \in F$ tal que para todo subgrupo $H < G$, se tenga que $\{\sigma(u) \mid \sigma \in H\}$ es una base de F sobre F^H.

5.2 Formas Bilineales sobre Campos Finitos

Dados un campo k y dos matrices simétricas $A, B \in M_n(k)$, tenemos un problema natural que es decidir si A y B son congruentes. Este problema lo resuelve completamente el Teorema de la Inercia de Sylvester (Teorema 4.52) para el caso en que $k = \mathbb{R}$. Este resultado también exhibe formas canónicas para matrices simétricas ante congruencia.

El caso de k algebraicamente cerrado y $\mathrm{caract}\, k \neq 2$, también es fácil; y se propuso como el Ejercicio 4.9.2. Cuando $k = \mathbb{Q}$, la situación es más complicada; pero se sabe la respuesta, que viene dada por el Teorema de Hasse-Minkowski, enunciado en la última sección de este capítulo.

En el caso de campos finitos, también hay una respuesta completa y sencilla al problema de congruencia, como veremos a continuación.

Atención: Desde ahora suponemos que k es finito, que $\mathrm{caract}\, k \neq 2$ y que V es un espacio vectorial de dimensión finita sobre k.

Teorema 5.5 *Dada una extensión de campos finitos F/k, el morfismo norma $N_k^F : F^\star \to k^\star$ es suprayectivo.*

Demostración: Sean $\circ(k) = q = p^m$, $[F : k] = n$ y $\mathrm{Gal}(F/k) = \langle\tau\rangle$, donde p es primo, $\tau = \sigma^m$ y σ es el morfismo de Frobenius. Entonces

$$a \in \ker N_k^F \Leftrightarrow a\tau(a)\cdots\tau^{n-1}(a) = 1;$$

pero $1 = a\tau(a)\cdots\tau^{n-1}(a) = a^{1+q+q^2+\cdots+q^{n-1}}$ tiene cuando más

$$1 + q + q^2 + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1}$$

soluciones, por lo que $\circ(\mathrm{Im}\, N_k^F) \geq q - 1$. Siendo clara la desigualdad $\circ(\mathrm{Im}\, N_k^F) \leq q - 1$, se tiene que $\mathrm{Im}\, N_k^F = k^\star$. □

Una forma bilineal B es **isotrópica** cuando existe un vector isotrópico ($v \neq 0$ con $B(v, v) = 0$). En caso contrario, la forma es **anisotrópica**.

Proposición 5.6 *Sea $B : V \times V \to k$ una forma simétrica bilineal isotrópica no degenerada. Entonces, dado $b \in k$, existe $v \in V$ tal que $B(v, v) = b$.*

Demostración: Supongamos que $0 \neq u \in V$ es isotrópico: $B(u, u) = 0$. Como B no es degenerada, existe $w \in V$ tal que $B(u, w) = 1$, entonces escribimos $v = cu + w$ y calculamos $B(v, v) = 2c + B(w, w)$, donde es posible resolver $c = [b - B(w, w)]/2$, para tener $B(v, v) = b$. □

Teorema 5.7 *Sea $B : V \times V \to k$ una forma simétrica bilineal no degenerada, tal que $\dim V \geq 2$. Entonces, dado $b \in k$, existe $v \in V$ tal que $B(v, v) = b$.*

Demostración: En vista del resultado anterior, podemos suponer que B es anisotrópica. Claramente, una vez que b sea dado, es suficiente encontrar un vector v tal que $B(v, v) = b$, en cualquier subespacio de V. Así, suponemos que $\dim V = 2$.

Como existe una base ortogonal de V, podemos también suponer que $B(v, v) = ax^2 + cy^2$ con $ac \neq 0$, si $v = (x, y)$. Dividiendo entre a, todavía podemos limitarnos al caso en que $B(v, v) = x^2 + cy^2$, pues si conseguimos que $x^2 + (c/a)y^2 = b/a$, entonces tendremos $ax^2 + cy^2 = b$.

La hipótesis de que B es anisotrópica significa que $x^2 + cy^2 \neq 0$ siempre que $(x, y) \neq 0$. Esto quiere decir que $-c$ no es un cuadrado en k. Así, la extensión $F = k(\sqrt{-c})/k$ es de grado dos. Lo que queremos demostrar es que la norma $N_k^F(x + y\sqrt{-c}) = x^2 + cy^2$ es suprayectiva; pero esto lo tenemos gracias al Teorema 5.5. □

Teorema 5.8 *Sea k un campo finito con $\operatorname{caract} k \neq 2$. Dada una forma simétrica bilineal no degenerada $B : V \times V \to k$, existe una base ortogonal $\{u_1, ..., u_n\}$ de V tal que $B(u_i, u_i) = 1$, siempre que $1 \leq i < n$; y $B(u_n, u_n) = d$, donde d es el discriminante de la forma.*

Demostración: Por el resultado anterior, si $\dim V = n \geq 2$, entonces existe $u_1 \in V$ tal que $B(u_1, u_1) = 1$.

Consideramos la descomposición $V = \langle u_1 \rangle \oplus \langle u_1 \rangle^\perp$. Si se tiene que $\dim \langle u_1 \rangle^\perp \geq 2$, entonces existe $u_2 \in \langle u_1 \rangle^\perp$ tal que $B(u_2, u_2) = 1$, y así sucesivamente, hasta que $\langle u_1, ..., u_{n-1} \rangle^\perp = \langle u_n \rangle$, que debe satisfacer $B(u_n, u_n) = d$. □

Teorema 5.9 *a) Dada una matriz simétrica invertible $A \in M_n(k)$, donde k es un campo finito con $\operatorname{caract} k \neq 2$, existe una matriz ortogonal P tal que $PAP^t = \operatorname{diag}(1, ..., 1, d)$, donde $d = \det A$.*

b) Las matrices $\operatorname{diag}(1, ..., 1, d_1)$ y $\operatorname{diag}(1, ..., 1, d_2)$ son congruentes si y sólo si $d_1 d_2^{-1} \in (k^\star)^2$.

Demostración: *a)* es la versión matricial del teorema anterior.

b) es consecuencia de que al discriminante lo consideramos un elemento de $k^\star/(k^\star)^2$. □

Observación. Hay campos infinitos F, donde el rango y el discriminante forman un conjunto completo de invariantes para sus formas cuadráticas. Lo que se necesita para que esto sea cierto, es poder resolver en F toda ecuación $ax^2 + by^2 = 1$ con $a, b \in F$ dados, como se aprecia al revisar la demostración del Teorema 5.8.

Esto ocurre por ejemplo, en el caso de una extensión algebraica infinita F de un campo finito k, pues si $a, b \in F$, entonces la extensión $k(a, b)/k$ es finita; y ahí podemos resolver esas ecuaciones.

Ejercicio

1. Demuestre que el número de clases de equivalencia ante congruencia para matrices simétricas $n \times n$ sobre un campo finito de orden impar es $2n + 1$.

5.3 La Densidad de Jacobson y sus Consecuencias

Sean U y V espacios vectoriales de dimensión no necesariamente finita, sobre un anillo de división D. En el conjunto $S = \mathrm{Hom}_D(U, V)$ definimos una topología decretando que los abiertos básicos sean de la siguiente forma: Fijamos $n \in \mathbb{N}$, n elementos linealmente independientes $u_1, ..., u_n \in U$; y n elementos arbitrarios $v_1, ..., v_n \in V$; entonces $\mathcal{O}(u_1, ..., u_n; v_1, ..., v_n) = \{A \in S \mid A(u_i) = v_i,\ \forall i\}$ es un abierto básico. Así, un conjunto abierto es una unión arbitraria de conjuntos $\mathcal{O}(u_1, ..., u_n; v_1, ..., v_n)$.

Esta topología hace de S un espacio de Hausdorff. Cuando la dimensión de U es finita, la topología es discreta. Nos interesa saber cuándo un subconjunto $S_0 \subseteq S$ es denso en S.

(1) Si $\dim U < \infty$, S_0 es denso en S exactamente cuando $S = S_0$.

(2) En general, S_0 es denso en S cuando S_0 es n-**transitivo** para todo $n \in \mathbb{N}$, es decir, cuando dados $\{u_1, ..., u_n\} \subseteq U$ linealmente independiente y elementos $v_1, ..., v_n \in V$, exista $L \in S_0$ tal que $L(u_i) = v_i$, para todo i.

Ejemplo. El conjunto $\mathrm{Hom}_D^F(U, V)$ de funciones lineales de rango finito, es siempre denso en $\mathrm{Hom}_D(U, V)$.

Sean U un grupo abeliano y R un subanillo del anillo de endomorfismos de U. Recordemos que un R-módulo $M \neq 0$ es simple o irreducible cuando sus únicos submódulos son M y 0. El Lema de Schur, p. 84, afirma que $D = \mathrm{End}_R U$ es un anillo de división, si U es irreducible como R-módulo.

Lema 5.10 *Sean $M \neq 0$ un R-módulo irreducible y $0 \neq a \in M$, entonces $Ra = M$.*

Demostración: Como Ra es un submódulo de M, es suficiente notar que $Ra \neq 0$. □

Teorema 5.11 (de la Densidad de Jacobson) *Sean U un grupo abeliano aditivo y R un anillo de endomorfismos de U. Supongamos que U es irreducible como R-módulo y que D es el anillo de división* $\mathrm{End}_R\, U$, *entonces R es denso en* $\mathrm{End}_D\, U$.

Demostración: El lema afirma que R es 1-transitivo. Sabiendo esto, demostraremos que
(*) *Dado $\{u_1, ..., u_n\} \subseteq U$ linealmente independiente, existe $r \in R$ tal que $ru_1 = \cdots = ru_{n-1} = 0$; con $ru_n \neq 0$.*

Procedemos por inducción en n, donde el caso $n = 1$ es la 1-transitividad de R. Así, suponemos que $n \geq 2$ y que **(*)** es cierta para $n - 1$.

Sea $I = \{r \in R \mid ru_1 = \cdots = ru_{n-2} = 0\}$. Este es un ideal izquierdo de R. La hipótesis inductiva afirma que $Iu_{n-1} \neq 0$; pero entonces existe $i \in I$ tal que $iu_{n-1} \neq 0$, por lo que $Iu_{n-1} \supseteq Riu_{n-1} = U$. Así, $Iu_{n-1} = U$.

Supongamos que **(*)** es falsa. Esto significa que para $i \in I$, se tiene

$$(iu_{n-1} = 0) \Rightarrow (iu_n = 0),$$

lo cual nos permite definir $T \in \mathrm{End}\, U$ como sigue: Para $u = iu_{n-1}$ con $i \in I$, escribimos $T(u) = iu_n$. Ahora afirmamos que $T \in \mathrm{End}_R\, U$.

Si $r \in R$ y $u \in U$ es tal que $u = ju_{n-1}$ con $j \in I$, entonces

$$rT(u) = rT(ju_{n-1}) = r(ju_n) = (rj)u_n = T(rju_{n-1}) = T(ru),$$

por lo que $T \in \mathrm{End}_R\, U$; y T resulta ser multiplicación por un escalar $\alpha \in D$. Por tanto, si $i \in I$, tenemos que $iu_n = T(iu_{n-1}) = i\alpha u_{n-1}$; de manera que I anula al vector $u_n - \alpha u_{n-1}$, que es linealmente independiente de $\{u_1, ..., u_{n-2}\}$, contradiciendo la hipótesis inductiva y demostrando **(*)**.

Sabiendo **(*)**, dados $\{a_1, ..., a_n\}$ linealmente independiente y $b_1, ..., b_n$ arbitrarios, existen $r_i \in R$ tales que $r_i a_j = \delta_{ij} b_j$, pues por ejemplo existe $r \in R$ tal que $ra_1 = \cdots = ra_{n-1} = 0,\ ra_n \neq 0$; y entonces existe $s \in R$ con $sra_n = b_n$, además de que $sra_1 = \cdots = sra_{n-1} = 0$.

Así, $(r_1 + \cdots + r_n)(a_i) = b_i$, para $i = 1, ..., n$. □

Se dice que un anillo satisface la condición descendente en ideales izquierdos cuando toda cadena estrictamente descendente de ideales izquierdos es finita. Esto es equivalente a exigir que todo conjunto no vacío de ideales izquierdos tenga un mínimo.

Ejemplo. Si D_n es el álgebra de matrices $n \times n$ sobre un anillo de división D, entonces es claro que D_n satisface la condición descendente en ideales izquierdos, pues estos son subespacios vectoriales. Es fácil ver que D_n es también un anillo simple.

Un R-módulo M es **fiel** cuando $rM = 0$ implica $r = 0$, para $r \in R$.

Teorema 5.12 (Wedderburn-Artin) *Sea R un anillo simple que satisface la condición descendente en ideales izquierdos, entonces $R \cong D_n$, para un anillo de división D.*

Demostración: Sea $U \neq 0$ un ideal izquierdo mínimo de R. Como el ideal $I = \{r \in R \mid rU = 0\}$ de R es bilateral, vemos que $I = 0$, por lo que U es un R-módulo fiel. Así, R es un anillo de endomorfismos de U.

El R-módulo U es irreducible, pues todo submódulo propio de U es un ideal izquierdo de R contenido propiamente en U.

Sea D el anillo de división $\mathrm{End}_R U$. Tenemos que R es denso en $\mathrm{End}_D U$.

Afirmamos que $\dim_D U$ es finita. Si $\mathcal{B} = \{u_1, u_2, ...\} \subseteq U$ es linealmente independiente sobre D, entonces los ideales izquierdos $I_m = \mathrm{an}(u_1, ..., u_m)$ de R forman una cadena estrictamente descendente, por la densidad de R en $\mathrm{End}_D U$. Concluimos que $\mathcal{B}$ es finito, que $\dim_D U$ es finita; y que $R = \mathrm{End}_D U \cong D_n$, para algún n. □

Corolario 5.13 *Sea A un álgebra simple de dimensión finita sobre un campo algebraicamente cerrado k. Entonces $A \cong k_n$ para algún n.*

Demostración: Como los ideales izquierdos de A son subespacios vectoriales, vemos que A satisface la condición descendente en ideales izquierdos. Entonces $A \cong D_n$ para algún n y un anillo de división D con $\dim_k D < \infty$, pues $\dim_k A < \infty$. Si $a \in D$, entonces $k(a)/k$ es una extensión finita de campos, por lo que $a \in k$. Así tenemos que $D = k$. □

Un subconjunto $S \subseteq \mathrm{End}_D U$ actúa **irreduciblemente** en U cuando los únicos subespacios W de U tales que $SW \subseteq W$ son 0 y U.

Lema 5.14 *Sea U un espacio vectorial de dimensión finita sobre un campo algebraicamente cerrado k y sea S un subálgebra de $\mathrm{End}_k U$ que actúa irreduciblemente. Entonces todo elemento de $\mathrm{End}_k U$ que conmuta con los elementos de S es un escalar.*

Demostración: Sea $T \in \mathrm{End}_k U$ tal que conmuta con los elementos de S y sea $\alpha \in k$ un valor característico de T. Entonces $T - \alpha$ también conmuta con los elementos de S.

Sea $W = \ker(T - \alpha)$. Afirmamos que W es invariante ante S: Si $u \in W$ y $L \in S$, entonces el cálculo

$$(T - \alpha)(Lu) = TLu - \alpha Lu = LTu - L\alpha u = L\alpha u - L\alpha u = 0$$

demuestra que $Lu \in W$. Así, $W = U$; y por tanto, $T = \alpha$. □

Teorema 5.15 (Burnside) *Sea A un álgebra de transformaciones lineales de un espacio vectorial U de dimensión finita sobre un campo algebraicamente cerrado k. Si A actúa irreduciblemente, entonces $A = \mathrm{End}_k U$.*

Demostración: Por el Teorema de la Densidad, $A = \mathrm{End}_D U$, donde D es el centralizador de $\mathrm{End}_A U$. Por el lema, $D = k$. □

Ejercicios

1. Sea V un espacio vectorial de dimensión finita sobre un anillo de división D. Demuestre que $\mathrm{End}_D V$ es un anillo simple que satisface la condición descendente en ideales izquierdos.

2. Sea R un anillo doblemente transitivo de transformaciones lineales en un espacio vectorial V sobre un anillo de división D. Demuestre que

   (a) R es denso en $\mathrm{End}_D V$.

   (b) $\mathrm{End}_R V = D$.

   (Sugerencia: Demuestre primero (b): Si existe alguna transformación $T \in (\mathrm{End}_R V) \setminus D$, entonces también existe $0 \neq v \in V$ tal que $\{v, Tv\}$ es linealmente independiente, por lo que podremos encontrar $S \in R$ con $Sv = 0$ y $STv = v$; pero entonces $0 \neq v = STv = TSv = 0$).

3. Demuestre las siguientes afirmaciones que incluyen un Teorema de Kolchin (b).

   a) Sean k un campo, $V \neq (0)$ un espacio vectorial de dimensión finita y $G < GL(V)$ que consiste de elementos unipotentes. Entonces existe un vector característico v común para los elementos de G, esto es $x(v) = v$, para todo $x \in G$.

   Un grupo, como G, que consiste de elementos unipotentes se llama unipotente.

   b) Todo grupo unipotente estabiliza una bandera y es triangulable.

   c) Todo grupo unipotente es nilpotente.

   Para demostrar a) se ofrecen las siguientes sugerencias:

   d) Redúzcase al caso en que k es algebraicamente cerrado, pues se trata de resolver sistemas de ecuaciones lineales.

   e) Redúzcase al caso en que V es un $k[G]$-módulo irreducible.

   f) Observe que para todos $x, y \in G$, se tiene $\dim V = \mathrm{tr}(xy) = \mathrm{tr}(y)$. De manera que para toda $n \in \mathrm{End}\, V$ con $x = 1+n$, para algún $x \in G$, se tiene que $\mathrm{tr}(y) + \mathrm{tr}(ny) = \mathrm{tr}(xy) = \mathrm{tr}(y)$ implica $\mathrm{tr}(ny) = 0$ para todo n de la forma anterior y todo $y \in G$.

   g) Invoque el Teorema de Burnside para obtener $\mathrm{tr}(ny) = 0$ para todo n de la forma anterior y todo $y \in \mathrm{End}\, V$. Deduzca que $n = 0$ y que $x = 1$.

5.4 Semisimplicidad

Fijemos un anillo asociativo R y consideremos módulos izquierdos sobre R.

Proposición 5.16 *Sea M un R-módulo. Las siguientes condiciones son equivalentes:*

1. M es suma de módulos simples.

2. M es suma directa de módulos simples.

3. Si E es un submódulo de M, entonces existe un submódulo P tal que $M = E \bigoplus P$.

Demostración: *1* $\Rightarrow$ *2*: Supongamos que $M = \sum_{i\in I} M_i$, con todo M_i simple. Sea J un subconjunto máximo de I tal que la suma $N = \sum_{j\in J} M_j$ sea directa. Afirmamos que $N = M$. Es suficiente ver que para todo $i \in I$, se tiene $M_i \subseteq N$; pero $M_i \cap N$ es 0 ó M_i. Como $M_i \cap N = 0$ implica que J no es máximo, se obtiene $M_i \cap N = M_i$, es decir, $M_i \subseteq N$.

2 $\Rightarrow$ *3*: Dado un submódulo E de M, consideramos a J, un subconjunto máximo de I tal que la suma $E + \coprod_{j\in J} M_j$ sea directa. El razonamiento anterior demuestra que $E + \coprod_{j\in J} M_j = M$.

3 $\Rightarrow$ *1*: Primero veremos que todo submódulo $N \neq 0$ de M contiene algún submódulo simple. A partir de un elemento $0 \neq a \in N$, consideramos al ideal izquierdo $I = \operatorname{an} a$ de R. Tenemos que $0 \neq Ra \cong R/I$ como R-módulos. Sea $\mathfrak{m}$ un ideal izquierdo máximo de R que contenga a I, entonces $\mathfrak{m}a$ es un submódulo máximo de Ra. Por hipótesis, existe un submódulo P tal que $M = \mathfrak{m}a \oplus P$, por lo que $Ra = \mathfrak{m}a \oplus (P \cap Ra)$. En estas condiciones, el módulo $P \cap Ra$ es simple y está contenido en N.

Sea E la suma de todos los submódulos simples de M. Afirmamos que $E = M$. Pues de no ser así, existiría un submódulo $L \neq 0$ con $M = E \oplus L$; pero entonces L contendría un submódulo simple ajeno a E. Esta contradicción concluye la demostración. □

Se dice que un R-módulo M es **semisimple** cuando satisface las condiciones de la proposición.

Proposición 5.17 *Todo submódulo y todo cociente de un módulo semisimple M, son semisimples.*

Demostración: Dado un submódulo N, sea E la suma de sus submódulos simples. Suponiendo $E \neq N$, existe un submódulo $L \neq 0$ de M tal que $M = E \oplus L$; pero entonces $N = E \oplus (L \cap N)$; y L contiene un submódulo simple de N ajeno a E. Esta contradicción demuestra la primera afirmación.

En cuanto a los cocientes, escribimos $M = N \oplus P$, donde sabemos que P es la suma de sus submódulos simples. Así, $M/N \cong P$ es semisimple. □

Se dice que un anillo R es **semisimple** cuando lo es como R-módulo. El siguiente resultado es inmediato.

Corolario 5.18 *Si R es un anillo semisimple, entonces todo R-módulo es semisimple.*

Se dice que un ideal izquierdo I de un anillo R es **simple** cuando lo es como R-módulo. Dos ideales izquierdos I, J de un anillo R se dicen **isomorfos** cuando lo son como R-módulos.

A partir de un anillo semisimple R, consideramos una colección de ideales izquierdos $\{E_i\}_{i\in I}$ que contenga exactamente un representante para cada clase de isomorfismo de ideales simples de R.

Teorema 5.19 *Sea R un anillo semisimple. Entonces:*

a) El número n de ideales simples de R no isomorfos entre sí es finito.

b) Si R_i es la suma de los ideales izquierdos de R isomorfos a E_i, cada R_i es un ideal bilateral de R y es un anillo cuyas operaciones son inducidas por las de R.

c) $R \cong \prod_i R_i$.

d) Si e_i es la identidad multiplicativa de R_i, para $i = 1, ..., n$; entonces $\{e_1, ..., e_n\}$ es una colección de idempotentes ortogonales con suma 1.

Demostración: Observemos primero que si M es un R-módulo simple, entonces para cada i se tiene que, ó bien $E_i \cong M$, ó bien $E_i M = 0$. Esto es cierto porque $E_i M$ es un submódulo de M; y $E_i M = M \Rightarrow E_i a = M$ para cualquier $a \in M$ tal que $E_i a \neq 0$, que a su vez da origen al isomorfismo $E_i \to M$ dado por $t \mapsto ta$.

Sea R_i la suma de los ideales izquierdos simples de R isomorfos con E_i. Es inmediato que $R = \sum_{i \in I} R_i$ y que $i \neq j \Rightarrow R_i R_j = 0$. En vista de esto y de que para cada i, R_i es un ideal izquierdo, se tiene que

$$R_i \subseteq R_i R \subseteq R_i R_i \subseteq R_i,$$

lo que demuestra que todo R_i es un ideal bilateral.

b) y *d)* Escribiendo $1 = e_1 + \cdots + e_n$, con $0 \neq e_i \in R_i$, para todo i, es inmediato que $e_i e_j = \delta_{ij} e_i$. Esto significa que $\{e_1, ..., e_n\}$ es una colección de idempotentes ortogonales con suma 1; y que e_i es la identidad multiplicativa del anillo R_i, cuyas operaciones son las inducidas por R.

a) Para cada $r \in R$ se tiene que $r = re_1 + \cdots + re_n$ y que $re_i \in R_i$, lo que demuestra que el conjunto de índices $I = \{1, ..., n\}$ es finito.

c) Como $R_i R_j = 0$ para $i \neq j$, tenemos que a suma $R = \sum_i R_i$ es directa y que $R = \prod_i R_i$. □

Teorema 5.20 *Sean R semisimple y $M \neq 0$ un R-módulo. Entonces*

$$M = \coprod_{i=1}^{n} R_i M,$$

con cada $R_i M$ la suma de los submódulos simples de M isomorfos con E_i.

Demostración: Por el Corolario 5.18, M es semisimple. Así sabemos que M es la suma de sus submódulos simples; por tanto, $M = RM = \coprod_i R_i M$, con cada $R_i M$ como en el enunciado. □

Teorema 5.21 *Sean k un anillo de división, V un espacio vectorial de dimensión finita sobre k y $R = \mathrm{End}_k V$. Entonces*

a) R es semisimple y todos sus módulos simples son isomorfismos.

b) V es un R-módulo simple.

Demostración: *b)* V es simple porque R actúa transitivamente en $V \setminus \{0\}$.

a) A partir de una base $\{u_1, ..., u_n\}$ de V, obtenemos un isomorfismo de R-módulos $\varphi : R \to V^n$ así: $\varphi(T) = (Tu_1, ..., Tu_n)$. Así, el R-módulo R es

una suma (directa) de R-módulos simples isomorfos y es semisimple. En la expresión $R \cong \prod_i R_i$ del Teorema 5.19, hay solamente un factor, porque R no tiene ideales bilaterales no triviales. □

Ejercicios

1. Sean V un espacio vectorial de dimensión finita sobre un campo algebraicamente cerrado k y $A \in \mathrm{End}_k V$. Demuestre que $k[A]$ es un anillo semisimple si y sólo si A es una transformación lineal semisimple.

2. Demuestre el Teorema de Maschke: Sean G un grupo de orden n y k un campo tal que $(\mathrm{caract}\, k) \nmid n$, entonces el álgebra de grupo $k[G]$ es semisimple.

   (Sugerencia: Demuestre que dados un $k[G]$-módulo E y un $k[G]$-submódulo F, entonces existe un $k[G]$-submódulo H de E tal que $E = F \oplus H$. A partir de la proyección lineal $\pi : E \to F$, considere la transformación dada por

   $$\varphi(x) = \frac{1}{n} \sum_{g \in G} g^{-1} \cdot (\pi(g \cdot x)).)$$

3. Demuestre que si R es un anillo que satisface la condición descendente en ideales izquierdos, entonces todo conjunto no vacío de ideales izquierdos contiene un elemento mínimo.

4. Demuestre que si R es un anillo que satisface la condición descendente en ideales izquierdos, entonces R es semisimple si y sólo si la intersección de todos sus ideales izquierdos máximos es cero.

5. Sea R un anillo conmutativo semisimple que satisface la condición descendente en ideales. Demuestre que R es isomorfo con un producto directo finito de campos.

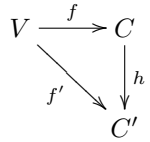
5.5 Algebras de Clifford

Sea V un espacio vectorial de dimensión n sobre un campo k con caract $\neq 2$. Supongamos que V tiene una forma cuadrática $q : V \to k$. Esto es,

$$q(v) = \sum_{i \leq j} a_{ij} v_i v_j, \text{ con } a_{ij} \in k, \text{ si } v = (v_1, ..., v_n).$$

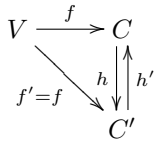
Observemos que entonces, la función $B : V \times V \to k$, definida como $B(x, y) = q(x + y) - q(x) - q(y)$ es una forma simétrica bilineal tal que $q(x) = \frac{1}{2} B(x, x)$.

El **álgebra de Clifford** $C = C(V) = C(V, q)$ de la forma q es un álgebra asociativo con unidad sobre k, junto con una función lineal $f : V \to C$ tal que $f(v)^2 = q(v)$ para todo $v \in V$; y que es universal con respecto a estas propiedades. Esto quiere decir que dados un álgebra asociativo C' y una función lineal $f' : V \to C'$ tal que $f'(v)^2 = q(v)$ para todo $v \in V$, existe un **único** morfismo de álgebras $h : C \to C'$ tal que $f' = h \circ f$, es decir, que hace conmutativo al diagrama



Unicidad: A partir de la definición, es claro que C es único, módulo isomorfismos de k-álgebras.

La imagen $f(V)$ genera a C: Sea C' el subálgebra generado por $f(V)$. Consideramos al morfismo inclusión $h' : C' \to C$, entonces existe un único morfismo $h : C \to C'$ que hace conmutativo al diagrama



pero $h' \circ h$ es la identidad en C, ya que $f = (h' \circ h) \circ f$, por lo que h' es suprayectivo y $C' = C$.

Existencia: Consideremos todas las parejas de álgebras asociativos C_i y funciones lineales $f_i : V \to C_i$, donde $f_i(V)$ genere a C_i y donde $f_i(v)^2 = q(v)$, para todo $v \in V$. Elijamos una pareja para cada clase de isomorfismo. Así obtenemos un conjunto índice I.

Sea $f : V \to \prod_{i \in I} C_i$ dado por $f(v)_i = f_i(v)$. Entonces el subálgebra C generado por $f(V)$ cumple la condición universal de la definición.

En el Problema 2 se exhibe un morfismo inyectivo de espacios vectoriales $\psi : V \to \mathcal{A}$, tal que $\psi^2(v) = q(v) \cdot 1$, para todo $v \in V$; y tal que $\mathcal{A}$ es un álgebra asociativo con uno.

Una **involución** es un automorfismo de orden dos.

Lema 5.22 *Existe una única involución i de C tal que $i[f(v)] = -f(v)$.*

Proposición 5.23 *Para cada $w^* \in V^*$, existe una i-derivación $d = d_{w^*} : C \to C$ tal que*

a) d es lineal.
b) $d(1) = 0$.
c) $d[f(v)] = w^*(v)$, para todo $v \in V$.
d) $d(cc') = d(c)c' + i(c)d(c')$, para todos $c, c' \in C$.

Demostración: Inventamos un símbolo nuevo u, para construir el álgebra de polinomios $C[u]$; pero definimos en ellos la multiplicación

$$(\sum_j a_j u^j)(\sum_r b_r u^r) = \sum_{j,r} a_j i^j(b_r) u^{j+r},$$

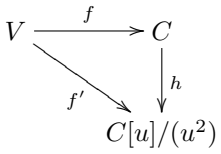
que es asociativa. De manera que tenemos un **álgebra de polinomios torcidos** $C[u]$, que da origen a los **números duales torcidos** $C[u]/(u^2)$, cuyos elementos son de forma $a + bu$ con $a, b \in C$.

Para cada $v \in V$, consideramos al elemento $f(v) + w^*(v)u \in C[u]$; y definimos una función lineal $f' : V \to C[u]/(u^2)$ así:

$$f'(v) = f(v) + w^*(v)u \pmod{u^2},$$

que satisface $f'(v)^2 = f(v)^2 + \{w^*(v)i[f(v)] + w^*(v)f(v)\}u = f(v)^2 = q(v)$.

Así, existe un único morfismo h que hace conmutativo al diagrama



Pero la proyección $h' : C[u]/(u^2) \to C$ es tal que $h' \circ h$ es la identidad en C, como puede verificarse en $f(V)$. Obtenemos que $h(c) = c + d(c)u$ para cierta función lineal $d : C \to C$ tal que $d[f(v)] = w^*(v)$. De la igualdad $h(cc') = h(c)h(c')$ se deducen las condiciones restantes $b)$ y $d)$. □

Teorema 5.24 *Sea $\{v_1, ..., v_n\}$ una base de V. Entonces el siguiente conjunto es una base de C:*

$$\{f(v_{i_1})f(v_{i_2})\cdots f(v_{i_r}) \mid 1 \le i_1 < i_2 < \cdots < i_r \le n;\ 0 \le r \le n\}.$$

Demostración: Como $f(V)$ genera a C, todo $c \in C$ es una combinación lineal de productos $f(v_{i_1})f(v_{i_2})\cdots f(v_{i_r})$.

Si $i > j$, entonces $f(v_i)f(v_j) + f(v_j)f(v_i) = f(v_i+v_j)^2 - f(v_i)^2 - f(v_j)^2 = 2B(v_i, v_j)$, por lo que $f(v_i)f(v_j) = 2B(v_i, v_j) - f(v_j)f(v_i)$; y todo $c \in C$ es una combinación lineal de elementos de nuestro conjunto, es decir, de aquellos $f(v_{i_1})f(v_{i_2})\cdots f(v_{i_r})$ con $i_1 < i_2 < \cdots < i_r$.

Si $\sum_{i_1<\cdots<i_r} a_{i_1\cdots i_r} f(v_{i_1})\cdots f(v_{i_r}) = 0$, entonces

$$0 = (d_{v^*_{i_r}} \circ \cdots \circ d_{v^*_{i_1}}) \left[\sum_{i_1<\cdots<i_r} a_{i_1\cdots i_r} f(v_{i_1})\cdots f(v_{i_r}) \right] = a_{i_1\cdots i_r};$$

y nuestro conjunto es linealmente independiente. □

Corolario 5.25 *El morfismo $f : V \to C$ es inyectivo.*

Así, de ahora en adelante, identificamos a V con $f(V)$.

Corolario 5.26 *Si* $\dim V = n$, *entonces* $\dim C = 2^n$.

Corolario 5.27 *Sean* V *y* V' *espacios vectoriales provistos de formas bilineales* B *y* B' *respectivamente. Si* $f : V \to V'$ *es una función lineal tal que* $B'(f(u), f(v)) = B(u, v)$ *para todos* $u, v \in V$, *entonces* f *se extiende a un único morfismo de álgebras* $f : C \to C'$. *Si* f *es inyectivo (suprayectivo, biyectivo o es un automorfismo), entonces su extensión a* C *también lo es. En particular, si* V *es un subespacio de* V', *entonces* $C(V)$ *es un subálgebra de* $C(V')$.

Ejemplos.

1. Cuando $q(v) = 0$ para todo $v \in V$, tenemos que $C(V) \cong \bigwedge V$. Si además $\dim V = 1$, entonces $C(V)$ consiste de elementos $a + bv$ con $a, b \in k$, donde $v^2 = 0$. Estos son los números duales.

2. Cuando $\dim V = 1$ y q es no degenerada, $C(V)$ tiene base $\{1, v\}$, con $v^2 = a \in k$. Distinguimos dos casos.

   Caso 1: $a \notin k^2$. Aquí, $C(V) = k(\sqrt{a})$.

   Caso 2: $a \in k^2$. Cambiando de base, se puede tener $v^2 = 1$. Usando $\operatorname{caract} k \neq 2$, vemos que $(1/2)(1 + v)$ y $(1/2)(1 - v)$ vienen a ser idempotentes ortogonales con suma uno, por lo que $C(V) = k((1/2)(1 + v)) \oplus k((1/2)(1 - v))$. Así, $C(V) \cong k \oplus k$. Cuando la característica es 2, entonces $\{1, 1+v\}$ es una base de $C(V)$ y $(1+v)^2 = 0$, de manera que $C(V)$ se identifica otra vez con los números duales.

3. Cuando $\dim V = 2$ y también $\operatorname{caract} k \neq 2$, entonces podemos escribir $q(xu + yv) = ax^2 + by^2$, para una base ortogonal $\{u, v\}$ de V con $q(u) = a, q(v) = b$. Supongamos que q es no degenerada, de manera que $ab \neq 0$. Aquí, $\{1, u, v, uv\}$ es una base de $C(V)$ con $uv = -vu$, que nos permite identificar a $C(V)$ con los **cuaternios generalizados**:

$$\alpha = p + qu + rv + suv, \quad \overline{\alpha} = p - qu - rv - suv$$
$$\alpha\overline{\alpha} = p^2 - aq^2 - br^2 + abs^2.$$

   El caso clásico ocurre cuando $a = b = -1$ y $k = \mathbb{R}$, siendo la forma $q(w) = -x^2 - y^2$ para $w = (x, y)$.

4. El caso clásico de álgebras de Clifford con $\dim V = n$, se da cuando $k = \mathbb{R}$ y q es negativa definida.

Se dice que un k-álgebra es **simple central** cuando no tiene ideales bilaterales y su centro es k. Si D es un anillo de división, D_n denota el álgebra de matrices $n \times n$ sobre D.

Proposición 5.28 *Sea V un espacio vectorial de dimensión dos, provisto de una forma cuadrática no degenerada. Entonces el álgebra de cuaternios generalizados $C(V)$ es simple central.*

Demostración: Sabemos que $C(V)$ tiene una base $\{1, u, v, uv\}$ que satisface $u^2 = a, v^2 = b$ y $uv = -vu$. Supongamos que $p, q, r, s \in k$ son tales que $\alpha = p + qu + rv + suv$ está en el centro de $C(V)$.

Calculamos: $u\alpha = qa + pu + sav + ruv,\ \alpha u = qa + pu - sav - ruv$. Esto implica $r = s = 0$. Después calculamos $v\alpha = pv - quv,\ \alpha v = pv + quv$. Esto implica $q = 0$. Así, $\alpha \in 1 \cdot k$.

Sea I un ideal bilateral de $C(V)$ y sea $\alpha = p+qu+rv+suv \in I$. Entonces $\beta = u\alpha + \alpha u = 2(qa + pu) \in I$ y también $v\beta + \beta v = 4qav \in I$, que a su vez implica que $4qav^2 = 4qab \in I$. Así, $q = 0$. De manera análoga, se puede obtener que $r = s = 0$, forzando entonces a que $p = 0$; y a que $\alpha = 0$. Así, $I = 0$. □

Proposición 5.29 *Sea V un espacio vectorial de dimensión dos, provisto de una forma cuadrática no degenerada. Entonces el álgebra de cuaternios generalizados $C(V)$ es un anillo de división o bien es isomorfo con el álgebra de matrices 2×2 sobre k.*

Demostración: Si la forma $\alpha\overline{\alpha} = p^2 - aq^2 - br^2 + abs^2$ del Ejemplo 3 representa a cero solamente de manera trivial, entonces $C(V)$ es un anillo de división:

$$\alpha \neq 0 \Rightarrow \alpha^{-1} = \frac{1}{\alpha\overline{\alpha}}\ \overline{\alpha}.$$

Si $C(V)$ no es un anillo de división, entonces como $C(V)$ es simple central y satisface la condición descendente en ideales izquierdos, por el Teorema de Wedderburn tenemos que $C(V) \cong D_m$, para un anillo de división D que es un álgebra sobre k. Como $\dim_k C(V) = 4$ y $C(V) \neq D$, se tiene que $D = k$ y que $m = 2$. Así, $C(V) \cong k_2$. □

Cuando el campo k es suficientemente grande, por ejemplo, cuando k es algebraicamente cerrado, siempre es posible conseguir que la forma $\alpha\overline{\alpha}$ del Ejemplo 3 represente a cero de manera no trivial, extrayendo una raíz cuadrada adecuada. Por lo que en este caso, $C(V) \cong k_2$.

Se dice que V es un **plano hiperbólico** cuando $\dim V = 2$ y V es isotrópico con una forma cuadrática no degenerada.

Proposición 5.30 *Todo plano hiperbólico V tiene una base $\{u, v\}$ tal que $q(u) = q(v) = 0$ y que $(u, v) = 1$.*

Demostración: Sea $0 \neq u \in V$ isotrópico, esto es, $q(u) = \frac{1}{2}(u, u) = 0$. Como el producto interno no es degenerado, existe $w \in V$ tal que $(u, w) = 1$, entonces el vector $v = -q(w)u + w$ satisface $q(v) = 0$ y $(u, v) = 1$. □

Ejemplo. El álgebra de Clifford $C(V)$ de un plano hiperbólico admite una base $\{1, u, v, uv\}$ tal que $u^2 = v^2 = 0$ y que $uv + vu = 1$, al encontrar u, v

isotrópicos con $(u,v) = \frac{1}{2}$; pero $\{u, v, uv, vu\}$ resulta ser una mejor base, ya que exhibe a $C(V)$ como el álgebra de matrices 2×2 sobre k así:

$$\begin{array}{cc} e_{11} = uv & e_{12} = u \\ e_{21} = v & e_{22} = vu \end{array}$$

Esto es porque $e_{11}e_{12} = uvu = u(1-uv) = u - u^2v = u = e_{12}$, etc.

Supongamos ahora que el espacio vectorial V tiene una base ortogonal $\{v_1, ..., v_n\}$; y que en $C(V)$ se tiene $v_i^2 = a_i \in k^\star$, para $1 \leq i \leq n$, de manera que $\prod_1^n a_i = \Delta$ es el discriminante de la forma.

Teorema 5.31 *Si n es par, entonces el centro Z de $C(V)$ es k. Si n es impar, entonces $dim_k\ Z = 2$ y Z está generado por $\{1, v_1v_2\cdots v_n\}$. En este último caso,*

$$Z \cong \begin{cases} k \oplus k, \ \textit{si } (-1)^{n(n-1)/2}\Delta \in k^2 \\ k(\sqrt{(-1)^{n(n-1)/2}\Delta}), \ \textit{en caso contrario} \end{cases}$$

Demostración: Observamos que conjugación con v_i deja fijo a cualquier monomio de $C(V)$, o lo envía a su negativo. Por tanto, Z es el espacio vectorial generado por los monomios de $C(V)$ que conmutan con todo v_i.

Al calcular $v_i(v_{i_1}v_{i_2}\cdots v_{i_r})v_i^{-1}$, el factor izquierdo v_i produce un número de cambios de signo al brincar hacia la derecha para ocupar su posición, mientras que el factor derecho v_i^{-1} produce otro número de cambios de signo al brincar hacia la izquierda hasta ocupar la suya.

Por tanto, cuando r es par, $v_{i_1}v_{i_2}\cdots v_{i_r}$ conmuta con todo v_i que no ocurre entre los v_{i_j}; y anticonmuta con los v_i que sí ocurren entre los v_{i_j}. Cuando r es impar, la situación es al revés. La conclusión es la del enunciado. □

Teorema 5.32 *Sea V un espacio vectorial con una forma q no degenerada tal que $n = \dim V \geq 3$ y sea U un subespacio de dimensión dos, donde la restricción q_0 de q es no degenerada. Escribimos $V = U \oplus U^\perp$; y q_1 la restricción de q a $U^\perp$. Entonces*

$$C(V,q) \cong C(U,q_0) \otimes C(U^\perp, -\delta q_1),$$

donde δ es el discriminante de q_0.

Demostración: La inyección canónica $U \hookrightarrow V$ da origen a una inyección $f : C(U,q_0) \hookrightarrow C(V,q)$.

Sea $\{u,v\}$ una base ortogonal de U, entonces $d = 2uv \in C(U,q_0)$ satisface $d^2 = -4q(u)q(v) = -B(u,u)B(v,v) = -\delta$. La inyección canónica $U^\perp \hookrightarrow V$ da origen a una inyección $C(U^\perp, q_1) \hookrightarrow C(V,q)$, que a su vez da origen a un morfismo $g : C(U^\perp, -\delta q_1) \hookrightarrow C(V,q)$, que envía a cualquier elemento $z \in U^\perp \subseteq C(U^\perp, -\delta q_1)$ al producto dz, pues $(dz)^2 = -\delta q(z)$.

Como $y(dz) = (dz)y$, para todos $y \in U,\ z \in U^{\perp}$, vemos que $f(U)$ centraliza a $g(U^{\perp})$. Así, gracias a la propiedad universal del producto tensorial, obtenemos un morfismo $\varphi : C(U, q_0) \otimes C(U^{\perp}, -\delta q_1) \to C(V, q)$ a partir de f y g. Aquí, $\varphi(y \otimes 1 + 1 \otimes z) = y + dz$, para $y \in U,\ z \in U^{\perp}$.

La igualdad $d^2 = -\delta$ demuestra que d es invertible en $C(U, q_0)$. Por otra parte, $yd = -dy$ en $C(U, q_0)$, para todo $y \in U$. Dados $y \in U,\ z \in U^{\perp}$, consideramos al elemento $w = (y \otimes 1) + (d^{-1} \otimes z) \in C(U, q_0) \otimes C(U^{\perp}, -\delta q_1)$; y calculamos:

$$w^2 = (q(y) \otimes 1) + (yd^{-1} + d^{-1}y) \otimes z + (1 \otimes q(z)) = q(y) + q(z) = q(y + z).$$

La propiedad universal de $C(V, q)$ produce el morfismo $\psi : C(V, q) \to C(U, q_0) \otimes C(U^{\perp}, -\delta q_1)$ tal que $\psi(y + z) = (y \otimes 1) + (d^{-1} \otimes z)$, para $y \in U,\ z \in U^{\perp}$.

Tenemos que $\varphi \circ \psi$ y $\psi \circ \varphi$ son los morfismos identidad en $C(V, q)$ y en $C(U, q_0) \otimes C(U^{\perp}, -\delta q_1)$ respectivamente, como se ve al evaluarlos en sus generadores. □

Dado un álgebra de Clifford $C(V)$, definimos $C^+(V)$ como el subálgebra de $C(V)$ generado por los elementos de grado par.

Proposición 5.33 *a) $C^+(V)$ es un álgebra de Clifford sobre un espacio de dimensión* $\dim V - 1$.

b) Si $\dim V = n$ *es impar, entonces $C(V) \cong C^+(V) \otimes Z$, donde Z es el centro de $C(V)$.*

Demostración: *a)*: Sea $\{u_1, ..., u_n\}$ una base ortogonal de V que satisface $u_1^2 = a_1, ..., u_n^2 = a_n$ en $C(V)$. Sean $v_1 = u_1u_n, ..., v_{n-1} = u_{n-1}u_n$, entonces $C^+(V)$ es el álgebra de Clifford del espacio $\langle v_1, ..., v_{n-1}\rangle$, donde v_i y v_j anticonmutan para $i \neq j$, con $v_1^2 = -a_1a_n, ..., v_{n-1}^2 = -a_{n-1}a_n$.

b): Esto es claro, pues $Z = \langle 1, u_1 \cdots u_n\rangle$, cuando n es impar. □

Teorema 5.34 *Sea q una forma cuadrática no degenerada sobre un espacio vectorial V de dimensión n sobre un campo k con* $\operatorname{caract} k \neq 2$. *Entonces:*

a) Si n es par, $C(V)$ es un producto tensorial de álgebras de cuaternios.

b) Si n es impar, $C(V)$ es un producto tensorial del centro con álgebras de cuaternios.

c) Si n es par y k es algebraicamente cerrado, $C(V)$ es isomorfo a un álgebra de matrices.

d) Si n es impar y k es algebraicamente cerrado, $C(V)$ es isomorfo a la suma directa de dos álgebras de matrices de la misma dimensión.

Demostración: Procedemos por inducción en n. Cuando $n = 1$, *b)* y *d)* se obtienen del Ejemplo 2; mientras que cuando $n = 2$, el Ejemplo 3 muestra que $C(V)$ es un álgebra de cuaternios; y la Proposición 5.29 afirma que $C(V)$ es un álgebra de matrices para k algebraicamente cerrado, por lo que se tienen *a)* y *c)*.

Supongamos que $n > 2$. Aquí podemos elegir un subespacio U de dimensión dos, donde la restricción q_0 de q sea no degenerada. Entonces $V = U \oplus U^\perp$ y la restricción q_1 de q a $U^\perp$ también es no degenerada. Por el Teorema 5.32, $C(V,q) \cong C(U,q_0) \otimes C(U^\perp, -\delta q_1)$, donde δ es el discriminante de q_0. Sabemos que $C(U,q_0)$ es un álgebra de cuaternios, por lo que la hipótesis inductiva aplicada a $C(U^\perp, -\delta q_1)$ nos permite obtener a). Si k es algebraicamente cerrado, c) es consecuencia de que un producto tensorial de álgebras de matrices también es un álgebra de matrices.

Ahora, b) es consecuencia de a) y de la Proposición 5.33; mientras que d) se obtiene de c) y de las Proposiciones 4.5.4 y 5.33. □

Ejercicios

1. Demuestre que si i es la involución del Lema 5.22 y si d es cualquier i-derivación de la Proposición 5.23, entonces $d \circ i + i \circ d = 0$ y $d^2 = 0$.

2. Sea V un espacio vectorial de dimensión finita sobre un campo k con $\operatorname{caract} k \neq 2$, provisto de una forma cuadrática $q : V \to k$. Sea $\mathcal{A} = \operatorname{End}_k(\bigwedge V)$.

   - Observe que $\mathcal{A}$ es un álgebra asociativo.
   - Para cada $v \in V$, defina $\ell_v : \bigwedge V \to \bigwedge V$ así: $\ell_v(w) = v \wedge w$, para todo $w \in \bigwedge V$. Observe que $\ell_v \in \mathcal{A}$ y que $\ell_v^2 = 0$.
   - Para cada $v \in V$, defina $g_v \in V^*$ así: $g_v(u) = B(v,u)$, para todo $u \in V$, donde B es la forma bilineal asociada a q. Sea d_v la antiderivación de la Proposición 5.23 correspondiente a g_v. Observe que $d_v \in \mathcal{A}$ y que $d_v^2 = 0$.
   - Para cada $v \in V$, defina $\psi(v) = \ell_v + d_v$. Observe que $\psi(v) \in \mathcal{A}$. Demuestre que $\psi(v)^2 = q(v) \cdot 1$ y que $\psi : V \to \mathcal{A}$ es inyectivo.

5.6 Teoremas de Frobenius y de Hurwitz

Concluimos el capítulo y el libro con la discusión de dos resultados clásicos, cuyas demostraciones utilisarán el material que hemos desarrollado en este capítulo.

Teorema 5.35 (Frobenius) *Sea $\mathcal{A}$ un álgebra de división de dimensión finita sobre $\mathbb{R}$. Entonces $\mathcal{A}$ es isomorfo con alguno de los siguientes objetos:*
Los números reales $\mathbb{R}$.
Los números complejos $\mathbb{C}$.
Los cuaternios reales $\mathbb{H}$.

Demostración: Paso 1: El conjunto $V = \{v \in \mathcal{A} \mid v^2 \leq 0\}$ es un subespacio vectorial de $\mathcal{A}$ de codimensión uno.

Si $\dim \mathcal{A} = m$, dado $a \in \mathcal{A}$, consideramos la multiplicación izquierda por a como la transformación lineal $T_a : \mathcal{A} \to \mathcal{A}$; así $T_a \in \operatorname{End} \mathcal{A}$. Como el morfismo $\mathcal{A} \to \operatorname{End}(\mathcal{A})$ dado por $a \mapsto T_a$ es inyectivo, identificamos a con su imagen T_a. El polinomio característico $p(X)$ de a se factoriza como

$$p(X) = (X - t_1) \cdots (X - t_r)(X - z_1)(X - \overline{z}_1) \cdots (X - z_s)(X - \overline{z}_s),$$

donde $m = r+2s$ y cada polinomio $X^2-2(\operatorname{Re} z_j)X+|z_j| = (X-z_j)(X-\overline{z}_j)$ es irreducible en $\mathbb{R}[X]$.

Dado que $p(a) = 0$ por Cayley-Hamilton y que $\mathcal{A}$ es un álgebra de división, obtenemos que o bien $a = t_i \in \mathbb{R}$ o bien $a^2 - 2(\operatorname{Re} z_j)a + |z_j| = 0$. En este último caso, escribimos $z = z_j$ para tener que el polinomio mínimo de a es $X^2 - 2(\operatorname{Re} z)X + |z|$ y que el polinomio característico de a es $(X^2 - 2(\operatorname{Re} z)X + |z|)^k$, donde $2k = m$.

Cuando $a \in V$, se da esta última situación; y además a es raíz de un polinomio cuadrático $X^2 + b$ con b real no negativo. De esto resulta que $a \in V$ si y sólo si $\operatorname{Re} z = 0$, es decir si y sólo si $\operatorname{tr} T_a = 0$, de donde se obtiene la afirmación.

Paso 2: Conclusión. Dado que $\mathcal{A} = \mathbb{R} \oplus V$, vemos que $\mathcal{A} = \mathbb{R}[V]$. Sea W un subespacio mínimo de V tal que $\mathcal{A} = \mathbb{R}[W]$.

Definimos una función simétrica bilineal en W así:

$$B(u, v) = -\frac{1}{2}(uv + vu).$$

Como $(u + v)^2 - u^2 - v^2 = uv + vu$, resulta que B tiene valores reales. Además, $B(u, u) > 0$ para $0 \neq u \in W$, pues $u^2 \leq 0$, de manera que B es positiva definida.

Sea $\{e_1, ..., e_n\}$ una base ortonormal de W. Esto significa que $e_i^2 = -1$ y que $e_i e_j = -e_j e_i$ para todos i, j. Así, el anillo de división $\mathcal{A}$ es isomorfo al álgebra de Clifford $C(W)$ asociada a una forma cuadrática negativa definida. Por lo tanto, resultan los siguientes casos:

- Cuando $n = 0$, se tiene que $\mathcal{A} \cong \mathbb{R}$.
- Cuando $n = 1$, se tiene que $\mathcal{A} \cong \mathbb{C}$.
- Cuando $n = 2$, se tiene que $\mathcal{A} \cong \mathbb{H}$.
- Cuando $n \geq 3$, el álgebra $C(W)$ nunca es un anillo de división:

  Sea $a = e_1 e_2 e_n$. Aquí, $a^2 = 1$, de manera que $(a + 1)(a - 1) = 1$ implicaría que $a = \pm 1$; y entonces $e_n = \pm(e_1 e_2)^{-1} = \pm e_2^{-1} e_1^{-1} = \pm(-e_2)(-e_1) = \mp e_1 e_2$, contradiciendo la independencia lineal de $\{e_1 e_2, e_n\}$. □

Teorema 5.36 (Hurwitz) *Sea V un espacio vectorial de dimensión n sobre un campo k de característica $\neq 2$, provisto de una forma cuadrática no degenerada $q : V \to k$ y de una multiplicación bilineal $u \cdot v$ tal que $q(u \cdot v) = q(u)q(v)$, para todos $u, v \in V$. Entonces $n = 1, 2, 4, 8$.*

Demostración: Paso 1: Es posible extender al campo k arbitrariamente, por lo que suponemos que k es algebraicamente cerrado.

Paso 2: Polarización de v en $q(uv) = q(u)q(v)$ produce

$$B(uv, uw) = q(u)B(v, w), \text{ para todos } u, v, w \in V,$$

donde la forma bilineal B es la asociada a q. Esto es porque se tienen

$$\begin{aligned} q(uv + uw) &= B(uv, uw) + q(uv) + q(uw) \\ q(u)q(v + w) &= q(u)[B(v, w) + q(v) + q(w)] \end{aligned}$$

Paso 3: Gracias al Paso 1, es posible elegir $u_1 \in V$ tal que $q(u_1) = 1$. Gracias al Paso 2, vemos que $\ell_{u_1} \in \operatorname{End} V$ es ortogonal y por tanto es invertible. Definimos una nueva multiplicación en V así:

$$u \circ v = \ell_{u_1}^{-1}(uv)$$

Aquí, $u_1 \circ v = v$, para todo $v \in V$, es decir, u_1 actúa como identidad izquierda. Escribiremos e en lugar de u_1, omitiremos el signo "$\circ$"; y usaremos esta nueva multiplicación, de manera que ahora tenemos una identidad izquierda $e \in V$ con $q(e) = 1$.

Paso 4: Conclusión. Consideramos al espacio $e^{\perp}$, de dimensión $n - 1$.

Sean $x \in e^{\perp};\ y, z \in V$. Le aplicamos q a la igualdad $(e + x)y = y + xy$, para obtener $q(e+x)q(y) = q(y)+q(xy)+B(y, xy)$; pero $q(e+x) = 1+q(x)$, por lo que obtenemos que $B(y, xy) = 0$, para todos $x \in e^{\perp},\ y \in V$.

Polarizamos en y esta última igualdad, para tener

$$0 = B(y + z, x(y + z)) = B(y, xy) + B(y, xz) + B(z, xy) + B(z, xz),$$

de manera que $B(y, xz)+B(z, xy) = 0$, para todos $x \in e^{\perp};\ y, z \in V$. Ahora reemplazamos z por xz, para tener $B(y, x(xz)) + B(xz, xy) = 0$, de donde obtenemos $B(y, x(xz)) + B(y, q(x)z) = 0$ ó bien $B(y, x(xz) + q(x)z) = 0$, para todos $y, z \in V$. Esto significa que $x(xz)+q(x)z = 0$, para todo $z \in V$.

Si ℓ_x es multiplicación izquierda por x, tenemos que $\ell_x^2 = -q(x)$, para todo $x \in e^{\perp}$. Por tanto, tenemos un morfismo del álgebra de Clifford $C = C(e^{\perp}, -q)$ hacia $\operatorname{End} V$, dado por $x \mapsto \ell_x$. Así, V es un C-módulo.

Supongamos n impar. Como k es algebraicamente cerrado, el Teorema 5.34*c*) afirma que C es un álgebra de matrices sobre un espacio W tal que $\dim W = 2^{(n-1)/2}$. Por el Teorema 5.21, C es un anillo semisimple y W es un C-módulo simple. Por el Teorema 5.20, V es una suma directa de copias de W. Por tanto, $n = r2^{(n-1)/2}$. Esto implica que $n = 1$.

Supongamos n par. Aquí, C es la suma de dos álgebras de matrices, cada una sobre un espacio de dimensión $2^{(n-2)/2}$, por el Teorema 5.34. Nuestra representación ahora es suma de r de copias del primer espacio más s copias del segundo. Por tanto, $n = (r + s)2^{(n-2)/2}$ y se tienen la

siguientes posibilidades:

$n = 2$	$r + s = 2$		
$n = 4$	$r + s = 2$		
$n = 6$	$r + s = \frac{3}{2}$	eliminada	
$n = 8$	$r + s = 1$		
$n > 8$	$n < 2^{(n-2)/2}$	eliminada	□

Observación. Aunque los álgebras de Clifford son asociativos, a la multiplicación bilineal $u \cdot v$ del Teorema de Hurwitz no se le exigió asociatividad, ni conmutatividad, ni existencia de identidad.

5.7 Ejercicios Generales

1. Un k-superálgebra es un k-álgebra A que admite una $\mathbb{Z}/2\mathbb{Z}$ graduación, esto es, una descomposición como suma directa de k espacios vectoriales $A = A_0 \oplus A_1$, tal que $A_i A_j \subseteq A_{i+j}$, donde los índices se suman módulo dos. Si $a \in A_i \setminus \{0\}$, escribimos $\mathrm{gr}(a) = i$.

   a) Defina los conceptos de morfimo graduado de superálgebras y de superideal.

   b) Demuestre que el núcleo de un morfimo graduado de superálgebras es un superideal y que el cociente de un superálgebra entre un superideal es un superálgebra.

2. Dadas dos k-superálgebras A y B, definimos el superproducto tensorial $A \widehat{\otimes} B$ como $[(A_0 \otimes B_0) \oplus (A_1 \otimes B_1)] \oplus [(A_0 \otimes B_1) \oplus (A_1 \otimes B_0)]$, con multiplicación $(a \otimes b)(c \otimes d) = (-1)^{\mathrm{gr}(b)\,\mathrm{gr}(c)} ac \otimes bd$.

   a) Demuestre que $A \widehat{\otimes} B$ es un superálgebra asociativo.

   Definimos el superconmutador $\{a, b\}$ de dos elementos homogéneos a y b como $ab - (-1)^{\mathrm{gr}(a)\,\mathrm{gr}(b)} ba$; a y b superconmutan cuando $\{a, b\} = 0$. Dos elementos arbitrarios c y d superconmutan cuando $\{c_i, d_j\} = 0$ para todas las componentes homogéneas c_i de c y d_j de d.

   b) Demuestre que dados dos morfismos graduados de superálgebras $i : A \to C$ y $j : B \to C$ tales que $i(A)$ superconmuta con $j(B)$, entonces tenemos un morfismo graduado inducido de superálgebras asociativos $A \widehat{\otimes} B \to C$.

3. Dado un espacio vectorial U provisto de una forma cuadrática q y con una descomposición ortogonal $U = U_1 \oplus U_2$, demuestre que existe un isomorfismo de superálgebras $C(U) \cong C(U_1) \widehat{\otimes} C(U_2)$.

4. Demuestre que $C(V \oplus V^*) \cong \mathrm{End}(\bigwedge V)$, donde $V \oplus V^*$ tiene la forma cuadrática standard, es decir, $B(u + u^*, v + v^*) = u^*(v) + v^*(u)$.

5.8 Enunciados

En esta sección enunciamos una serie de resultados importantes.

Teorema 1 (Burnside) *Si p y q son números primos, entonces todo grupo de orden $p^a q^b$ es soluble.*

Teorema 2 (Feit-Thompson) *Todo grupo finito cuyo orden sea impar, es soluble.*

Teorema 3 (Nielsen-Schreier) *Todo subgrupo de un grupo libre, es libre.*

Teorema 4 (Kronecker-Weber) *Sea K una extensión abeliana finita de $\mathbb{Q}$. Entonces existe una raíz de la unidad ζ tal que $K \subseteq \mathbb{Q}[\zeta]$.*

Teorema 5 (Cartan-Dieudonné) *Si V es un espacio vectorial provisto de un producto interno y $\dim V = n$, entonces toda transformación ortogonal es un producto de cuando más n simetrías.*

Teorema 6 (Dirichlet) *Sean m y n enteros primos relativos. Entonces el conjunto de primos p tales que $p \equiv m (\mathrm{mod}\, n)$ es infinito.*

Teorema 7 (Lindemann-Weierstrass) *a) Si $\alpha_1, ..., \alpha_n$ son números algebraicos, linealmente independientes sobre $\mathbb{Q}$, entonces $\{e^{\alpha_1}, ..., e^{\alpha_n}\}$ es algebraicamente independiente sobre $\overline{\mathbb{Q}}$.*

b) Si $\alpha_1, ..., \alpha_n$ son números algebraicos distintos, entonces $\{e^{\alpha_1}, ..., e^{\alpha_n}\}$ es linealmente independiente sobre $\overline{\mathbb{Q}}$.

Teorema 8 (Hasse-Minkowski) *Dos matrices cuadradas y simétricas $A, B \in M_n(\mathbb{Q})$ son congruentes sobre $\mathbb{Q}$ si y sólo si lo son sobre $\mathbb{R}$ y sobre los campos p-ádicos $\mathbb{Q}_p$ para todo primo p.*

Errata de la versión anterior

- p.14: La condición 3 es superflua.
- p.16 línea - 6: Sea $A = \{(x_1, ..., x_p) \in G^p \mid x_1 \cdots x_p = 1\}$.
- p.30 Ejercicio 3: Sean k un campo arbitrario y $G = GL_n(k)$ el grupo multiplicativo de las matrices invertibles $n \times n$ con coeficientes en k. Definimos al grupo $B = \{[a_{ij}] \in G \mid a_{ij} = 0 \text{ cuando } i > j\}$, así como al grupo $U = \{[a_{ij}] \in B \mid a_{ii} = 1 \text{ para toda } i\}$. Demuestre que:

  a) $U \lhd B < G$.

  b) $U = (B, B)$, suponiendo que $n \geq 2$ y que $\circ(k) \geq 4$.

  c) U es nilpotente.

  d) B es soluble.

  e) $B \cong U \ltimes T$.
- p.40 línea 10:

$$Zpq = \langle a \mid a^{pq} = 1 \rangle = \langle a, b \mid a^p = 1, b^q = 1, aba^{-1}b^{-1} = 1 \rangle$$

- p.42 Teorema 1.77: Sea A un grupo abeliano tal que ...
- p.56 Problema 7:

$$e_i e_j = \delta_{ij} e_i, \ e_1 + \cdots + e_n = 1.$$

- p.63 línea 11: Recíprocamente, sea $\mathfrak{p}$ un ideal primo de R tal que $\mathfrak{p} \cap S = \varnothing$. Entonces $\varphi^{-1}(S^{-1}\mathfrak{p}) = \mathfrak{p}$, pues $(\mathfrak{p} : s) = \mathfrak{p}$ para todo $s \in S$.
- p.63 línea 15: Así, $S^{-1}\mathfrak{p}$ es primo.
- p.71 línea 7: Usando la norma N de $\mathbb{Q}[\sqrt{-19}]$ ...
- p.74 línea - 13: $\mathfrak{c}(f) = \text{m.c.d.}\{a_0, a_1, ..., a_n\}$.
- p.83 línea 4: Concluimos que $R(f, g) = S$. $\square$

- p.87 línea 7: ... es nilpotente. Supondremos que nuestros anillos son conmutativos.

- p.103 línea - 14: b) K es el campo generado por $\cup_\sigma \sigma(F)$, ...

- p.112 línea - 15: Por el Lema 3.33, tenemos que $[F:k] \leq \circ(\mathcal{G})$; pero $[F:k] = \circ(\mathrm{Gal}(F/k))$ y $\mathcal{G} \subseteq \mathrm{Gal}(F/k)$ implican que $\mathcal{G} = \mathrm{Gal}(F/k)$ y que F/k es una extensión finita.

- p.115

  Consideremos la cadena $K = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{n-1} = F$, donde $K_i = K_{i-1}(T_i)$, de manera que K_i/K_{i-1} es simple con elemento primitivo T_i, que es raíz del polinomio

$$f_i(X) = \frac{f(X)}{(X - T_1)\cdots(X - T_{i-1})} \in K_{i-1}[X],$$

  que es de grado $n - i + 1$. Esto implica que $[K_i : K_{i-1}] \leq n - i + 1$, para $1 \leq i \leq n-1$. Observe que $f_i(X) \in k[s_1, ..., s_n, T_1, ..., T_{i-1}][X]$. Como

$$\prod_{i=1}^{n-1}(n - i + 1) = n(n-1)\cdots 2 = n! = \circ(S_n) = [F : K],$$

  se ve que $[K_i : K_{i-1}] = n - i + 1$, para $1 \leq i \leq n-1$; y que $f_i(X) = \mathrm{Polmin}(T_i, K_{i-1})$.

  Dado un polinomio arbitrario $p(T_1, ..., T_n) \in R$, podemos usar la relación $T_n = s_1 - T_1 - \cdots - T_{n-1}$, para eliminar T_n en favor de las otras T_i y de s_1. Después usamos $f_{n-1}(X)$, que es de grado dos, para eliminar a T_{n-1}^2, expresándolo como polinomio en T_{n-1} de grado ≤ 1 con coeficientes en $k[s_1, ..., s_n, T_1, ..., T_{n-2}]$. Continuamos este proceso hasta expresar a $p(T_1, ..., T_n)$ como combinación lineal de los $n!$ monomios $T_1^{r_1}T_2^{r_2}\cdots T_n^{r_n}$, con $r_i \leq n - i$, para $1 \leq i \leq n$, con coeficientes en $k[s_1, ..., s_n]$. Además, para cada i, el conjunto $\{1, T_i, ..., T_i^{n-i}\}$ es una base de K_i sobre K_{i-1}, por lo que los $n!$ monomios mencionados forman una base de F sobre K. Esto implica que la expresión de $p(T_1, ..., T_n)$ indicada, es única. Observe que en ninguno de estos monomios aparece T_n, por lo que si $p(T_1, ..., T_n)$ es simétrico, entonces solamente aparece el término constante.

- p.137 línea 7: Sea F/k una extensión finita de Galois ...

- p.138 El penúltimo párrafo debe decir así:

  La cadena de campos correspondiente: $k = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_s = F$, es tal que cada extensión F_{i+1}/F_i es de Galois con grupo cíclico de orden p_i, por lo que F_1/k es de tipo 1), 2) ó 3). El tipo 2) sólo puede

ocurrir cuando k contenga a las raíces p_1-ésimas de la unidad; y la demostración concluye por inducción. En caso contrario, adjuntamos una raíz m-ésima primitiva de la unidad ζ al campo k, donde $m = \prod_{p_i \neq \text{caract}} p_i$.

- p.150 La demostración del Teorema 3.78 debe iniciar así:

  Podemos suponer que $n \geq 3$. Gracias a los Teoremas 3.48 y 3.36 que afirman que existen extensiones de $\mathbb{F}_p$ de cualquier grado para cualquier primo p; y que estas son simples, puede verse que existen $f_1(X) \in \mathbb{F}_2[X]$ mónico, irreducible de grado n, $g_2(X) \in \mathbb{F}_3[X]$ mónico, irreducible de grado $n-1$; así como $f_3(X) \in \mathbb{F}_5[X]$ mónico, de grado n, con factorización irreducible consistente en un polinomio cuadrático, junto con uno o dos factores de grado impar. Definimos $f_2(X) = X g_2(X)$ y observamos que es posible tener todo $f_i(X)$ separable.

- p.164 línea 10: generado por las expresiones $a \otimes b + b \otimes a$ para todos $a \in M,\ b \in N$;

Bibliografía

[1] Artin, E. **Geometric Algebra,** John Wiley, 1988.

[2] Artin, M. **Algebra,** second edition, Birkhäuser, 2003.

[3] Atiyah, M. F. and MacDonald, I. G. **Introduction to Commutative Algebra,** Perseus Books, 1994.

[4] Bourbaki, N. **Éléments de mathématique. Algèbre,** Chapitres 1 à 3, Masson-Dunod, 1970; Chapitres 4 à 7, Masson-Dunod, 1981. **Algebra 1,** Springer, 1989; **Algebra 2,** Springer, 1990.

[5] Burnside, W. **Theory of Groups of Finite Order,** second edition, Dover, 1955.

[6] Childs, L. **A Concrete Introduction to Higher Algebra,** Undergraduate Texts in Mathematics, second edition, Springer, 1995.

[7] Cox, D. A., Little, J. B. and O'Shea, D. **Ideals, Varieties and Algorithms,** second edition, Springer, 1997.

[8] Jacobson, N. **Basic Algebra I,** second edition, W. H. Freeman, 1989; **Basic Algebra II,** W. H. Freeman, 1974.

[9] Kempf, G. **Algebraic Structures,** Friedrich Vieweg, 2003.

[10] Lang, S. **Algebra,** third revised edition, Graduate Texts in Mathematics, Springer, 2002.

[11] Lidl, R. and Niederreiter, H. **Introduction to Finite Fields and Their Applications,** Cambridge University Press, 1986.

[12] van der Waerden, B. L. **Algebra,** Springer, 2003.

Para cubrir los requisitos de Algebra Lineal y para continuar con ese tema:

[13] Birkhoff, G. and MacLane, S. **A Survey of Modern Algebra,** fourth edition, Macmillan, 1977.

[14] Curtis, C. W. **Linear Algebra. An Introductory Approach,** fourth edition, seventh printing, Undergraduate Texts in Mathematics, Springer, 1999.

[15] Fraleigh, J. B. and Beauregard, R. A. **Linear Algebra,** Addison-Wesley, 1990.

[16] Greub, W. **Linear Algebra,** fourth edition, Graduate Texts in Mathematics, Springer, 1975.

[17] Greub, W. **Multilinear Algebra,** second edition, Springer, 1978.

[18] Halmos, P. R. **Finite-Dimensional Vector Spaces,** second edition, fifth printing, Undergraduate Texts in Mathematics, Springer, 1993.

[19] Hoffman, K. and Kunze, R. **Linear Algebra,** second edition, Prentice-Hall, 1971.

[20] Kaplansky, I. **Linear Algebra and Geometry, a second course,** Dover, 1995.

[21] Lang, S. **Linear Algebra,** third edition, Undergraduate Texts in Mathematics, Springer, 1987.

[22] Nering, E. D. **Linear Algebra and Matrix Theory,** second edition, John Wiley, 1995.

[23] Noble, B. and Daniel, J. W. **Applied Linear Algebra,** third edition, Prentice-Hall, 1988.

Para encontrar más ejercicios de grado de dificultad más variado:

[24] Herstein, I. N. **Topics in Algebra,** second edition, Wiley, New York 1975.

[25] Lang, S. **Undergraduate Algebra,** third edition, Undergraduate Texts in Mathematics, Springer, 2005.

Indice Alfabético